

Tilburg University

Bodies and background

Van den Stock, J.B.

Publication date:
2010

Document Version
Publisher's PDF, also known as Version of record

[Link to publication in Tilburg University Research Portal](#)

Citation for published version (APA):
Van den Stock, J. B. (2010). *Bodies and background: Contextual influences on perception of emotional body language*. [s.n.].

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Bodies and background:
Contextual influences on perception of emotional body
language

Proefschrift ter verkrijging van de graad van doctor aan de Universiteit van Tilburg, op gezag van de rector magnificus, prof. Philip Eijlander, in het openbaar te verdedigen ten overstaan van een door het college voor promoties aangewezen commissie in de aula van de Universiteit op vrijdag 4 juni 2010 om 10:15 uur door Van den Stock Jan geboren op 25 februari 1978 te Aalst (België).

Promotor: professor doctor Beatrice de Gelder

Contents

Introduction.....	7
References	14
Chapter 1: Real faces, real emotions: perceiving facial expressions in naturalistic contexts of voices, bodies and scenes.....	17
Introduction	17
Making space for context effects in models of face perception.....	18
Face perception and categorization.....	19
Similarities between facial expressions and other affective signals in perceptual and neurofunctional processes	19
From a face module to a face processing network	20
Face processing includes subcortical and cortical areas	22
Body context effects on facial expressions.....	23
Perception of facial expression is influenced by the bodily expressions	23
Possible mechanisms underlying body context effect	25
Facial and bodily expressions in the context of scenes	29
Facial expressions in the context of the affective prosody of voices	32
Summary and conclusions.....	34
References	36
Chapter 2: Standing up for the body. Recent progress in uncovering the networks involved in processing bodies and bodily expressions	43
Introduction	43
Similarities in perceptual routines: configuration and feature based perception of faces, bodies and other visual objects	44
The neurophysiology of body perception. Implications for neural representation and time course of processing.....	45
Similarities and differences in neurofunctional basis of faces and bodies	52
Emotional modulation of body selective processing.....	70
Static vs. Dynamic body perception.....	71
Bodies may be processed without attention and visual awareness	73
Learning from lesions	75
Models of body perception.....	77
Conclusion.....	79
References	80
Chapter 3: Seeing bodily expressions in context: contagion between bodily expression and scene.....	89

Introduction	89
Method	89
Results and Discussion.....	93
References	101
Chapter 4: Body expressions influence recognition of facial expressions and emotional prosody	103
Introduction	103
Experiment 1: recognition of bodily expressions	105
Method	105
Results	106
Discussion	107
Experiment 2: the influence of body expressions on recognition of facial expressions.....	108
Method	109
Results	109
Discussion	110
Experiment 3: The influence of body language on recognition of voice prosody	111
Method	112
Results	113
Discussion	114
General discussion	115
References	117
Chapter 5: Human and animal sounds influence recognition of body language	121
Introduction	121
Method	123
Results.....	126
Discussion	129
References	132
Chapter 6: Instrumental music influences recognition of emotional body language	135
Introduction	135
Method	136
Results.....	138
Discussion	140
References	144
Chapter 7: Moving and being moved. The relative importance of dynamical information for residual face processing in clinical populations and brain damaged patients	147

Introduction	147
Face perception: some antecedents	148
The brain basis of face perception in neurologically intact individuals: Perceiving movement.....	149
Neurophysiological studies in monkeys	154
Visual object agnosia and face agnosia or prosopagnosia	154
The neural correlates of face deficits.....	155
Still vs. dynamic face images in patient studies.....	156
Conclusion.....	158
References	159
Chapter 8: Prosopagnosia	167
Definitions and concepts	167
Inability to process the face as a whole.....	169
Acquired and developmental prosopagnosia	169
Face-specificity of prosopagnosia.....	170
Behavioural and neural correlates of prosopagnosia.....	172
References	173
Chapter 9: Neural correlates of perceiving emotional faces and bodies in developmental prosopagnosia: an event-related fMRI-study.....	177
Introduction	177
Methods.....	182
Results.....	187
Discussion	197
References	202
Chapter 10: Huntington's disease impairs recognition of angry and instrumental body language	209
Introduction	209
Method	210
Results.....	213
Discussion	215
References	218
Chapter 11: Perceiving emotions from bodily expressions and multisensory integration of emotion cues in schizophrenia.....	221
Introduction	221
Experiment 1: Recognition of static body language	222
Methods	222
Results	224

Discussion	225
Experiment 2: Multisensory integration of dynamic body language and vocalizations.....	226
Methods	227
Results	227
Discussion	229
General discussion	230
References	233
Discussion.....	237
The social brain hypothesis	237
Emotions are expressed in the whole body	238
faces convey information about identity	238
bodies convey information about actions	239
Emotional expressions are prone to contextual influences	240
time course of crosscategorical affective influence.....	241
neuro-anatomy of crosscategorical influences.....	242
References	246
Acknowledgements	249

Introduction

Humans are considered to be among the most social organisms on the planet. Indeed, the days we do not come in contact with conspecifics are exceptional. Humans normally take part in a wide range of heterogeneous social interactions on a daily basis. Therefore, the adaptive value of interpreting the intentions and emotions of others is considered to be significant.

Although the importance of social functioning and perception of the affective states of others is now generally acknowledged, research on emotion perception has only taken off in the last decades. The focus of emotion perception research has been primarily on how we perceive, process and recognize facial expressions. The findings so far have led to formulation of theories and models that do or do not capitalize on a ‘special’ status for faces in the perceptual system (Bruce & Young, 1986; de Gelder, Frissen, Barton, & Hadjikhani, 2003; Haxby, Hoffman, & Gobbini, 2000; Kanwisher, McDermott, & Chun, 1997).

In our natural environment however, faces are not perceived in isolation and usually co-occur with a wide variety of visual, auditory, olfactory, somatosensory and gustatory stimuli. The main stimulus categories that are naturally associated with faces include whole bodies, voices and background scenes. This thesis moves beyond facial expression perception and focuses on how perception of emotional body language is shaped partially by contextual stimuli like natural scenes and vocal expressions. The content of this thesis can be divided into two sections: section one (chapters one to six) describes findings from research on normal subjects and section two (chapters seven to eleven) focuses on investigations in clinical populations.

Respecting the scientific chronology and the importance of facial expression research issues for the research agenda on bodily expressions, chapter one provides a review of studies addressing contextual influences on the perception of facial expressions. The emphasis of this review lies on how bodies, scenes and voices influence our perception of faces and facial expressions, both at the behavioral and the neural level. Chapter two also provides a review, but shifts the focus from contextual influences on face perception to neural correlates of

perceiving bodies and bodily expressions. These two review chapters provide an illustrative ‘contextual background’ and introduction to the following chapters that report empirical studies investigating diverse contextual influences on body perception.

The empirical chapters report studies using different methods to investigate the respective hypotheses. In the behavioral experiments, we focused on psychophysics and cognitive aspects of emotion processing. We have aimed to use paradigms that are most optimal in the operationalisation of the research question. In all psychophysical behavioral tasks, we aimed to keep response alternatives to a minimum in order to minimize the cognitive load of the task. Furthermore, verbal labeling was avoided when possible in order to tap primarily into the target processes of emotion perception and to minimize language effects. Our investigations of multisensory integration were based on experimental conflict situations and the crossmodal bias paradigm (de Gelder & Bertelson, 2003). In this paradigm, auditory emotional expressions are simultaneously presented with visual stimuli that display an emotional expression that is either congruent or incongruent with the auditory stimulus. Participants are instructed to categorize the emotion expressed by one of both modalities and to ignore the other modality. For example, fearful and happy vocal expressions are simultaneously presented with fearful or happy body language, while the task instruction states to categorize the emotion expressed by the body expression and to ignore the information provided by the voice. The difference between the congruent and incongruent conditions gives an indication of multisensory integration processes.

Data from behavioral psychophysical experiments reveal information about behavioral processes but are less suitable to investigate the time course or stage in the perceptual process where the observed effects occur, for example whether the nature of observed effects is early perceptual rather than late more cognitive. Carefully designed behavioral experiments allow inferences about the nature of perceptual effects (for example making use of recalibration or adaption paradigm), but the optimal methodology to investigate timing issues requires the use of techniques with a high temporal resolution, like event related potentials (ERP) or magnetoencephalogram (MEG). These techniques allow recording of electrical signals produced by the brain in the order of milliseconds. ERP data have low spatial resolution on the other hand and are therefore limited regarding information about the brain regions involved in emotional processes.

To investigate the functional neuro-anatomy of emotion processes, we made use of functional magnetic resonance imaging (fMRI). This technique requires that subjects perform a task while the magnetic characteristics of regional blood flow in the brain (the so-called blood oxygenation level dependent (BOLD) signal) are being recorded by an MRI-scanner. fMRI measurements reveal information about the amount of oxygenated and de-oxygenated blood levels throughout the brain during different experimental conditions. Using the subtraction method, this technique enables to make inferences about which brain areas require more or less oxygenated blood supply during one experimental condition, compared to another. fMRI thus provides a useful method to investigate neural networks involved in perception and cognition.

The fMRI data reported in chapter three focuses on brain regions involved in perception of bodily expressions embedded in neutral or threatening scenes in normal subjects. This provides information about brain areas that process both bodies and scenes. However, data from normal subjects is limited concerning which brain regions are either necessary or sufficient in performing the experimental task. This can be investigated with experiments that include patients with brain damage. Patient data are reported in chapters nine (prosopagnosia), ten (Huntington's disease) and eleven (schizophrenia). In these studies, we aimed to establish single dissociations, meaning that a patient population (associated with a characteristic brain abnormality) is impaired at performing task X, but not task Y. This is for example the case in chapter ten: patients with Huntington's disease are significantly worse than normal control subjects in recognizing angry and instrumental whole body expressions, whereas there is no significant difference between both groups for fearful and sad body expressions. This indicates that successful recognition of angry and instrumental body expressions is more dependent on the affected brain regions in patients with Huntington's disease (primarily caudate nucleus and putamen) than perceiving fearful and sad bodies. Ideally, patient studies aim to report double dissociations. This implies that patients with lesions in brain area X are able to perform task A, but impaired at performing task B, whereas patients with lesions in brain area Y are able to perform task B, but impaired at task A. This would demonstrate that brain region Y is critical for performing task A and brain region X is necessary for performing task B.

Combining imaging techniques and patient research allows investigation of more subtle research questions, for example what brain areas are recruited to perform a task when brain regions on which successful task performance relies are damaged (Vandenbulcke, Peeters, Van Hecke, & Vandenberghe, 2005).

The following provides an overview of chapters three to eleven, including the empirical goals and predictions where applicable.

Chapter three describes the study investigating the neural correlates of perceiving whole body expressions embedded in emotionally congruent or incongruent scenes. The idea for performing this study was built on recent findings of contextual influences of emotional scenes on facial expression processing, revealed by behavioral and ERP data (Righart & de Gelder, 2006, 2008a, 2008b). Combining these findings with models of body perception (de Gelder, 2006) and scene perception (Bar, 2004) led to the formulation of the research question and experimental setup making use of the experimental conflict situation in order to investigate how the brain integrates and processes emotional information from different sources, in this case bodies and scenes. One of the aims of this study was to explore the brain regions that are sensitive to (in)congruence between body expressions and scenes. Our predictions included activation of regions involved in emotion perception like the amygdala and regions involved in conflict processing like the anterior cingulate gyrus. However, the results showed that the integration of both stimulus categories mainly takes place in the known scene perception and body perception areas, namely the extrastriate body area (Downing, Jiang, Shuman, & Kanwisher, 2001), the parahippocampal place area (Epstein & Kanwisher, 1998) and the retrosplenial complex (Bar & Aminoff, 2003). The only brain area that was specifically sensitive to body x scene interaction corresponded to the lateral occipital complex (Malach et al., 1995). The latter finding is also reported by a recent study also using bodies embedded in scenes (Peelen, Fei-Fei, & Kastner, 2009).

Chapter four reports three experiments on how we recognize whole body expressions and the influence of body expressions on the perception of facial expressions (cfr supra) and vocal expressions. This study was set up in order to explore behavioral similarities between face and body perception. Previous studies have shown that fearful faces are the most difficult to

recognize (Ekman, 1976) and we hypothesized that this would also be the case for bodily expressions. In the first experiment participants performed a body expression matching task and we predicted lower accuracies and higher reaction times for fearful expressions compared to the other expressions. The results indicate good recognition of all emotions, with fear indeed being the hardest to recognize. In the second experiment, categorizations of the expression in the face of a face-body compound stimulus were strongly influenced by the bodily expression, compatible with a previous study using a similar paradigm but different emotions (Meeren, van Heijnsbergen, & de Gelder, 2005). Moreover, we observed that the influence of the body on the facial expression was a function of the ambiguity of the facial expression. In the third experiment, we focused on crossmodal effects using the experimental conflict situation. This experiment was carried out to investigate whether the previously reported influence of facial expressions on the recognition of voice prosody (de Gelder & Vroomen, 2000) extends to bodily expressions. We hypothesized an influence of the emotion conveyed by the task irrelevant bodily expression on the recognition of the emotional tone of voice. The findings indeed show a bias in the recognition of voice prosody towards the emotion expressed by the body expression. The results described in chapter four illustrate the importance of emotional whole body expressions in communication either when viewed on their own or, as is often the case in realistic circumstances, in combination with facial expressions and emotional voices.

Building on the results of experiment 3 in chapter four, we designed two audiovisual experiments described in chapters five and six. The experimental setup was similar to that of experiment 3 in chapter four, but now the task was reversed and focused on the visual modality. Subjects were instructed to categorize the emotion expressed by the body expression and to ignore the auditory information. We constructed dynamic face-blurred bodily expressions displaying a person engaged in an everyday action (grasping a glass) in a realistic situation (a living room). These visual stimuli were combined with human and animal vocalizations (chapter five) and instrumental music (chapter six). The main goal of these experiments was to investigate whether affective crossmodal bias effects also apply for audiovisual stimuli of which both modalities are unmistakably produced by a different source. Our prediction was that a crossmodal bias effect would also obtain for these stimulus combinations and this was confirmed by the data as the results indicated that recognition of

body language is biased towards the emotion expressed by the simultaneously presented auditory information, whether it consists of human or of animal sounds or instrumental music.

Chapter seven is the first chapter of the clinical section. Chapter seven provides a review of perception of dynamic facial expressions in clinical populations. It forms a bridge to the remaining chapters that focus on whole body perception in clinical populations. In chapter eight we review and discuss the notion of prosopagnosia, a peculiar symptom that relates to face recognition deficits. This is a prelude to chapter nine that describes how we used event-related fMRI to investigate how neutral and emotional facial and bodily expressions are processed in the brain of patients suffering from prosopagnosia. This experiment was set-up in the rationale that investigation of face perception deficits in prosopagnosia requires the use of a refined control stimulus category. Since on the one hand faces contain information of many dimensions (for example identity, emotion, gender, age...) and on the other hand our perceptual system has been bombarded with faces from birth, very few stimulus categories are suitable to serve as face-control. The stimulus category that is probably the most comparable to faces on a wide range of attributes as well as on the perceptual expertise dimension is bodies. Therefore bodies seem to be the preferable control category. The second consideration in setting up this study was the fact that there was up to that moment no available data on the neural correlates of facial expression perception in developmental prosopagnosics (subjects with prosopagnosia but without neurological antecedents). Our predictions were based on a study with acquired prosopagnosia (prosopagnosia following brain damage) that showed how emotional faces activated brain regions associated with emotion perception, while areas involved in face perception were damaged (de Gelder et al., 2003). The results were surprisingly in accordance with the notion of dissociable processing systems of facial emotion and facial identity: we observed normal activation for emotional faces and reduced activation for neutral faces. This finding was later replicated by a different research group (Peelen, Lucas, Mayer, & Vuilleumier, 2009). The second main finding of this study was that, compared to control subjects, faces trigger more activation in body areas and bodies trigger more activation in face areas in developmental prosopagnosics.

In chapter ten, we describe our research into the perception of whole body expressions in patients with Huntington's disease. Patients with HD exhibit motor impairments as well as cognitive and emotional deficits. The reason to investigate body language recognition in HD

was twofold. First, the growing body of evidence linking motor production and motor perception, as illustrated by research on mirror neurons (Rizzolatti & Craighero, 2004) brings forth the hypothesis of a link between motor production (impairment) and emotional motor perception (impairment). Second, the alleged importance of subcortical structures in emotional processes (de Gelder, 2006; Panksepp, 1998) consequently leads to the logical consideration that damage to subcortical structures disrupts emotion perception. Previously, impairments in the ability to recognize emotional stimuli have mostly been investigated by using facial expressions and emotional voices. Chapter ten reports how we tested recognition of emotional body language (instrumental, angry, fearful and sad) in a group of HD patients and their matched controls. The results indicate that HD patients are impaired in recognizing both instrumental and angry whole body postures. Furthermore, the body language perception deficits are correlated with measures of motor deficit. The results therefore suggest a close relationship between emotion recognition (specifically anger) and motor abilities.

The rationale behind the experiments reported in chapter eleven join with that of chapter four. The aim of the study in chapter eleven was to extend findings from facial expression research in schizophrenia to whole body expressions. The chapter reports how these patients perceive static whole body expressions but also if crossmodal bias effects previously reported for faces (de Jong, Hodiament, Van den Stock, & de Gelder, 2009) are replicable with bodies (cfr chapter five). As in previous chapters and based on the many similarities between faces and bodies, the predictions were in parallel with what is known from face research in schizophrenia. The results of Experiment 1 show emotion recognition impairment in the schizophrenic group and the findings of Experiment 2 show that schizophrenics are more influenced by the auditory information than controls, but only when the auditory information consists of human vocalizations. These findings indicate that emotion recognition difficulties in schizophrenia extend to whole body expressions and multisensory integration, particularly in naturalistic audiovisual combinations.

References

- Bar, M. (2004). Visual objects in context. *Nat Rev Neurosci*, 5(8), 617-629.
- Bar, M., & Aminoff, E. (2003). Cortical analysis of visual context. *Neuron*, 38, 347-358.
- Bruce, V., & Young, A. (1986). Understanding face recognition. *Br J Psychol*, 77 (Pt 3), 305-327.
- de Gelder, B. (2006). Towards the neurobiology of emotional body language. *Nature Reviews Neuroscience*, 7(3), 242-249.
- de Gelder, B., & Bertelson, P. (2003). Multisensory integration, perception and ecological validity. *Trends Cogn Sci*, 7(10), 460-467.
- de Gelder, B., Frissen, I., Barton, J., & Hadjikhani, N. (2003). A modulatory role for facial expressions in prosopagnosia. *Proc Natl Acad Sci U S A*, 100(22), 13105-13110.
- de Gelder, B., & Vroomen, J. (2000). The perception of emotions by ear and by eye. *Cognition and Emotion*, 14(3), 289-311.
- de Jong, J. J., Hodiamont, P. P., Van den Stock, J., & de Gelder, B. (2009). Audiovisual emotion recognition in schizophrenia: reduced integration of facial and vocal affect. *Schizophrenia Research*, 107(2-3), 286-293.
- Downing, P. E., Jiang, Y., Shuman, M., & Kanwisher, N. (2001). A cortical area selective for visual processing of the human body. *Science*, 293(5539), 2470-2473.
- Ekman, P., & Friesen, W. V. (1976). Pictures of Facial Affect. *Palo Alto, CA: Consulting Psychologists Press*.
- Epstein, R., & Kanwisher, N. (1998). A cortical representation of the local visual environment. *Nature*, 392(6676), 598-601.
- Haxby, J. V., Hoffman, E. A., & Gobbini, M. I. (2000). The distributed human neural system for face perception. *Trends Cogn Sci*, 4(6), 223-233.
- Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: a module in human extrastriate cortex specialized for face perception. *J Neurosci*, 17(11), 4302-4311.
- Malach, R., Reppas, J. B., Benson, R. R., Kwong, K. K., Jiang, H., Kennedy, W. A., et al. (1995). Object-related activity revealed by functional magnetic resonance imaging in human occipital cortex. *Proc Natl Acad Sci U S A*, 92(18), 8135-8139.
- Meeren, H. K., van Heijnsbergen, C. C., & de Gelder, B. (2005). Rapid perceptual integration of facial expression and emotional body language. *Proc Natl Acad Sci U S A*, 102(45), 16518-16523.

- Panksepp, J. (1998). *Affective neuroscience: The foundation of human and animal emotions*. New York: Oxford University Press.
- Peelen, M. V., Fei-Fei, L., & Kastner, S. (2009). Neural mechanisms of rapid natural scene categorization in human visual cortex. *Nature*, 460(7251), 94-97.
- Peelen, M. V., Lucas, N., Mayer, E., & Vuilleumier, P. (2009). Emotional attention in acquired prosopagnosia. *Soc Cogn Affect Neurosci*, 4(3), 268-277.
- Righart, R., & de Gelder, B. (2006). Context influences early perceptual analysis of faces--an electrophysiological study. *Cereb Cortex*, 16(9), 1249-1257.
- Righart, R., & de Gelder, B. (2008a). Rapid influence of emotional scenes on encoding of facial expressions: an ERP study. *Social cognitive and affective neuroscience*, 3(3), 270-278.
- Righart, R., & de Gelder, B. (2008b). Recognition of facial expressions is influenced by emotional scene gist. *Cognitive, affective & behavioral neuroscience*, 8(3), 264-272.
- Rizzolatti, G., & Craighero, L. (2004). The mirror-neuron system. *Annu Rev Neurosci*, 27, 169-192.
- Vandenbulcke, M., Peeters, R., Van Hecke, P., & Vandenberghe, R. (2005). Anterior temporal laterality in primary progressive aphasia shifts to the right. *Ann Neurol*, 58(3), 362-370.

Chapter 1: Real faces, real emotions: perceiving facial expressions in naturalistic contexts of voices, bodies and scenes¹

Introduction

For a while “Headless Body in Topless Bar” counted as one of the funniest lines to have appeared in US newspapers. But headless bodies and bodiless heads figure only in crime catalogues and police reports and are not part of our daily experience, at the very least not part of the daily experience that constitutes the normal learning environment in which we acquire our face and body perception expertise. Yet, except for a few isolated studies, the literature on face recognition has not yet addressed the issue of context effects in face perception. By ‘context’ we mean here the whole naturalistic environment that is almost always present when we encounter a face.

Why has context received so little attention and what, if any, changes would we need to make to mainstream models of face and facial expression processing if indeed different kinds of context have an impact on how the brain deals with faces and facial expressions? Discussions on context influences and their consequences for how we read and react to an emotion from the face have a long history (Fernberger, 1928). But the kind of context effects that were investigated in the early days would nowadays qualify as so called late effects or post-perceptual effects, related as they are to the overall (verbal) appraisal of a stimulus rather than to its online processing. In contrast, the context effects we have specifically targeted in recent studies are those that are to be found at the perceptual stage of face processing.

In this chapter we review recent investigations of three familiar naturalistic contexts in which facial expressions are frequently encountered: whole bodies, natural scenes and emotional voices (See also Ambady and Weisbuch, this volume) . In the first section we briefly review recent evidence that shifts the emphasis from a categorical model of face processing, based on the assumption that faces are processed as a distinct object category with their dedicated perceptual and neurofunctional basis, towards more distributed models where different

¹ This chapter is based on de Gelder, B., & Van den Stock, J. (in press). Real faces, real emotions: perceiving facial expressions in naturalistic contexts of voices, bodies and scenes. In Calder A., Haxby, J., Johnson, M., & Rhodes, G. (Eds.) *The Handbook of Face Perception* . Oxford University Press.

aspects of faces (like direction of gaze and emotional expression) are processed by different brain areas and different perceptual routines and show how these models are better suited to represent face perception and face-context effects. In the second section we look in detail at one kind of context effect, as found in investigations of interactions between facial and bodily expressions. We sketch a perspective in which context plays a crucial role, even for highly automated processes like the ones underlying recognition of facial expressions. Some recent evidence of context effects also has implications for current theories of face perception and its deficits.

Making space for context effects in models of face perception

Older theories on face perception have tended to restrict scientific investigations of face perception to issues of face vs. object categorization. The major sources of evidence for category specificity of face perception are findings about its temporal processing windows and neurofunctional basis. But this debate is not settled and recent evidence now indicates that the temporal and spatial neural markers of face categorization are also sensitive to some other non-face stimuli (for a review of such overlap between spatial and temporal markers of face and body specificity, see B. de Gelder et al., 2009). Furthermore, it is becoming increasingly clear that the presence of an emotional expression influences even those relatively early and relatively specific neural markers of category specificity like the N170 and the face area in fusiform gyrus. Finally, distributed models as opposed to categorical models of face processing seem more appropriate to represent the relation between face perception, facial expression perception and perceptual context effects as they represent the various functional aspects of facial information and allow for multiple entry points of context into ongoing face processing. Finally, models must also include the role of subcortical structures shown to be important components of face and facial expression processes.

Face perception and categorization

Much of the face recognition literature has been dominated by the view that face processing proceeds at its own pace, immune to the surrounding context in which the face is encountered. In line with this, one of the major questions in the field continues to be that of the perceptual and neurofunctional bases of faces. An important assumption has been and continues to be that faces occupy a neurofunctional niche on their own, such that face representations co-exists with but does not overlap with object representations, a view that in one sense or another is linked to the notion of modularity. Typical characteristics of modular processing as viewed in the eighties and brought to a broad audience by Fodor (1983) are mainly that processing is mandatory, automatic and insulated from context effects. What was originally a theoretical argument purporting to separate syntactic from the more intractable semantic aspects of mental processes became for a while the focus of studies using brain imaging (Kanwisher, McDermott, & Chun, 1997). A research program fully focused on category specificity is unlikely to pay attention to perceptual context effects on face processing. In contrast, more recent distributed models of face processing appear more suited to accommodate the novel context findings (B. de Gelder, Frissen, Barton, & Hadjikhani, 2003; Haxby, Hoffman, & Gobbini, 2000).

Similarities between facial expressions and other affective signals in perceptual and neurofunctional processes

Seeing bodily expressions is an important part of everyday perception and scientific study of how we perceive whole body expressions has taken off in the last decade. Issues and questions that have been addressed in face research are also on the foreground in research on whole body expressions (see B. de Gelder et al., 2009 for a review). This is not surprising, considering the fact that faces and bodies appear together in daily experience. It may be not so surprising that perception of faces and bodies show several similarities at the behavioural and neuro-functional level. For example, both faces and bodies are processed configurally, meaning as a single perceptual entity, rather than as an assemblage of features. This is reflected in the perceptual processes triggered when face and body stimuli are presented upside-down (the inversion effect): recognition of faces and bodies presented upside-down is relatively more impaired than recognition of inverted objects, like houses (Reed, Stone, Bozova, & Tanaka, 2003). Also, a comparison of perception of upright and inverted faces reveals that the time course of the underlying brain mechanisms is similar for upright and

inverted bodies (Stekelenburg & de Gelder, 2004). The presence of a bodily expression of fear in the neglected field also significantly reduces attention deficits in neurological populations (Tamietto, Geminiani, Genero, & de Gelder, 2007), just as has been reported for faces (Vuilleumier & Schwartz, 2001). As will be shown in detail in the later sections, perception of bodily expressions activates some brain areas that are associated with the perception of faces (for reviews, see B. de Gelder, 2006; Peelen & Downing, 2007. See also section II).

From a face module to a face processing network

Categorical models of face processing (e.g. Kanwisher *et al.*, 1997) tend to assume that the core of face processing consists of a dedicated brain area or module that is functionally identified by contrasting faces with a small number of other object categories mostly by using passive viewing conditions. All other dimensions of face processing corresponding to other dimensions of face information (emotion, age, attractiveness, gender...) are viewed as subsequent modulations of the basic face processing ability implemented in the brain's face area(s). In contrast, distributed models for face perception also consider other aspects of faces besides person identity (Adolphs, 2002; Adolphs, Damasio, Tranel, Cooper, & Damasio, 2000; B. de Gelder *et al.*, 2003; B. de Gelder & Rouw, 2000; Haxby *et al.*, 2000; Haxby *et al.*, 1994; Haxby *et al.*, 1996; Hoffman & Haxby, 2000; Puce, Allison, Asgari, Gore, & McCarthy, 1996). In distributed models, different areas of the brain process different attributes of the face, such as identity (FFA and the occipital face area (OFA), gaze direction (superior temporal sulcus (STS) and expression and/or emotion analysis (OFC, amygdala, anterior cingulate cortex, premotor cortex, somatosensory cortex).

Clinical cases constitute critical tests for theoretical models, and patients suffering from a deficit in face recognition or prosopagnosia (Bodamer, 1947) have long served as touchstone for models of face processing (see also chapters by Young, Calder, and Kanwisher and Barton). Available fMRI studies targeting face perception in prosopagnosics so far show inconsistent results (see Van den Stock, van de Riet, Righart, & de Gelder, 2008 for an overview), but very few of those studies included facial expressions or compared emotional with neutral faces (see Calder, this volume). Configural processing as measured by the inversion effect is a hallmark of intact face processing skills and a few studies have reported that the normal pattern of the inversion effect does not obtain when a face perception disorder

is present whether of acquired or of developmental origin (B. de Gelder & Rouw, 2000; but see McKone & Yovel, 2009). We investigated whether adding an emotional expression would normalize their face processing style with respect to the inversion effect. We presented neutral and emotional faces to patients with acquired prosopagnosia (face recognition deficits following brain damage) with lesions in FFA, inferior occipital gyrus (IOG) or both. Our study showed that emotional but not neutral faces elicited activity in other face related brain areas like STS and amygdala and, most importantly, that most of these patients showed a normal inversion effect for emotional faces as well as normal configural processing as measured by in a part-to-whole face identity matching task when the faces were not neutral but expressed an emotion (B. de Gelder et al., 2003). In a follow up fMRI study with patients suffering from developmental prosopagnosia (prosopagnosia without neurological history), we presented neutral and emotional (fearful and happy) faces and bodies and the results showed normal activation in FFA for emotional faces (fearful and happy) but lower activation for neutral faces, compared to controls (Van den Stock, van de Riet et al., 2008) (see Figure 1).

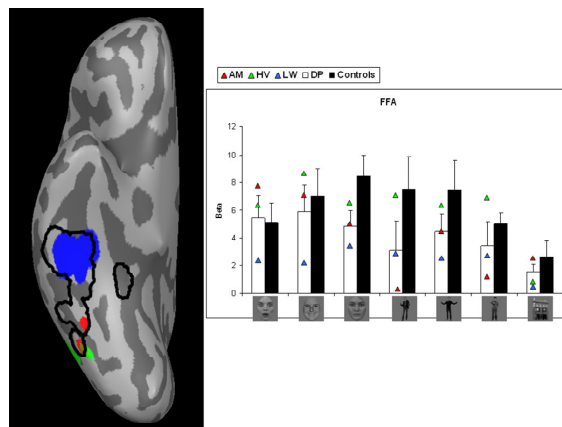


Figure 1. Face-specific BOLD-activation in right fusiform face area (FFA) when comparing faces (fearful/happy/neutral) with houses. Left: Areas are shown on an inflated right hemisphere. Activation maps of the control subjects are collapsed and displayed by the black contours. Activation of the individual developmental prosopagnosics (DPs) is plotted in color. Right: beta-values in the plotted areas. Conditions represent from left to right: fearful faces, happy faces, neutral faces, fearful bodies, happy bodies, neutral bodies and houses. White columns display the average value of three prosopagnosics, black columns show the average value of the controls. Triangles represent the individual values of the DPs. Error bars represent one standard error of the mean (SEM). (Van den Stock et al., 2008b)

Increased activation for emotional faces compared to neutral faces in FFA has since been reported in an acquired prosopagnosia case by others also (Peelen, Lucas, Mayer, & Vuilleumier, 2009).

Electrophysiological studies are crucial for investigating distributed face models because the limited time resolution of fMRI does not allow one to conclude that all dimensions of facial information indeed necessarily depend on activity in the fusiform face area. Studies using electroencephalogram (EEG) or magnetoencephalogram (MEG) data initially provided support for face modularity, in the sense that there appeared to be a unique time window for a stimulus to enter the face processing system. EEG and MEG investigations into face perception have characterised two early markers in the temporal dynamics of face perception: a positive waveform around 100ms (P1) and a negative waveform around 170ms (N170) after stimulus onset indicating the time course of dedicated brain mechanisms sensitive to face perception. It is a matter of debate where in the brain these waveforms originate, whether in early extrastriate areas, STS or fusiform gyrus (FG) and what type of processing mechanism these waveforms reflect, whether global encoding, object categorization or configural processing (see B. de Gelder et al., 2006 for a review).

Face processing includes subcortical and cortical areas

Finally, we have shown, as have other groups, that patients with striate cortex damage can process and recognize faces presented in their blind visual field (Andino, Menendez, Khateb, Landis, & Pegna, 2009; B de Gelder & Tamietto, 2007; B. de Gelder, Vroomen, Pourtois, & Weiskrantz, 1999; Morris, de Gelder, Weiskrantz, & Dolan, 2001; Pegna, Khateb, Lazeyras, & Seghier, 2005) and for which they have no conscious perception. For this and other reasons not relevant here, the involvement of subcortical structures in face perception also needs to be represented in a distributed model of face processing as we sketched in de Gelder et al (2003). Masking studies performed with neurologically intact observers, on residual visual abilities for faces and facial expressions in cortically blind patients and on face processing skills of infants with immature visual cortex converge to provide tentative evidence for the importance of subcortical structures. Research indicates that the distributed brain network for face perception encompasses two main processing streams: a subcortical pathway from superior colliculus and pulvinar to the amygdala that is involved in rudimentary and mostly nonconscious processing of salient stimuli like facial expressions (B. de Gelder, Pourtois, van

Raamsdonk, Vroomen, & Weiskrantz, 2001; B. de Gelder et al., 2008; B. de Gelder, Vroomen et al., 1999; Morris et al., 2001; Morris, Ohman, & Dolan, 1998; Pegna et al., 2005) and a more familiar cortical route from the lateral geniculate nucleus (LGN) via primary visual cortex to OFA, FFA and STS, sub serving fine grained analysis of conscious perception. Feed forward and feedback loops, especially between amygdala and striate cortex, OFA, FFA and STS (Amaral & Price, 1984; Carmichael & Price, 1995; Catani, Jones, Donato, & Ffytche, 2003; Idaka et al., 2001; Morris, Friston et al., 1998; Vuilleumier, Richardson, Armony, Driver, & Dolan, 2004) support the interaction between these routes to contribute ultimately to a unified and conscious percept (but see Cowey, 2004; Pessoa, McKenna, Gutierrez, & Ungerleider, 2002).

In summary, clinical phenomena like prosopagnosia and affective blindsight form an important contribution to the current understanding of face perception. Distributed face processing models that neuro-anatomically include subcortical structures and incorporate the many dimensions of faces like emotional expression appear to resonate best with the empirical data.

Body context effects on facial expressions

Of all the concurrent sources of affective signals that routinely accompany our sight of a facial expression, the body is by far the most obvious and immediate one. We review recent evidence for this perceptual effect and follow with a discussion of possible mechanisms underlying body context effects.

Perception of facial expression is influenced by the bodily expressions

Research on the simultaneous perception of faces and bodies is still sparse. Two behavioural studies directly investigated how our recognition of facial expressions is influenced by accompanying whole body expressions (Meeren, van Heijnsbergen, & de Gelder, 2005; Van den Stock, Righart, & de Gelder, 2007). Meeren et al. (2005) combined angry and fearful facial expressions with angry and fearful whole body expressions to create both congruent

(fearful face on fearful body and angry face on angry body) and incongruent (fearful face on angry body and angry face on fearful body) realistically looking compound stimuli (see Figure 2). These were briefly (200ms) presented one by one while the participants were instructed to categorize the emotion expressed by the face and ignore the body. The results showed that recognition of the facial expression was biased towards the emotion expressed by the body language, as reflected by both the accuracy and reaction time data. In a follow-up study, facial expressions that were morphed on a continuum between happy and fearful were once combined with a happy and once with a fearful whole body expression (Van den Stock *et al.*, 2007). The resulting compound stimuli were presented one by one for 150ms, while the participants were instructed to categorize the emotion expressed by the face in a 2 alternative forced choice paradigm (fear or happiness). Again, the ratings of the facial expressions were influenced towards the emotion expressed by the body and this influence was highest for facial expressions that were most ambiguous (expressions that occupied an intermediate position on the morph continuum). Evidence from EEG-recordings during the experiment shows that the brain responds to the emotional face-body incongruency as early as 115ms post stimulus onset (Meeren *et al.*, 2005). The reverse issue, whether perception of bodily expressions is influenced by facial expression has not been studied so far. However, natural synergies between facial and bodily expressions predict emotional spill over between the face and the body as exists between the facial expression and the voice (B. de Gelder & Bertelson, 2003).



Figure 2. Stimulus examples of congruent and incongruent face-body compounds (Meeren et al., 2005)

Possible mechanisms underlying body context effect

A few different explanations are suggested by body context effect. First, one may view these effects as providing support for a thesis that has a long history in research on facial expressions and states that facial expressions seen on their own are inherently ambiguous (Frijda, 1986). A different approach may be that emotions are intimately linked to action preparation and that action information is provided much more specifically by bodily than by facial expressions. A third consideration is that there may be considerable overlap between the neurofunctional basis of facial and bodily expressions such that showing either the face or the body also automatically triggers representation of the other.

Facial expressions may be inherently ambiguous

Does the strong impact of bodily expressions on judging facial expressions provide evidence for drawing the more radical conclusion that judgments of facial expressions are entirely context sensitive? Some recent studies have indeed suggested so. Adopting our methodology Aviezer et al. (2008) used disgust pictures with an average recognition of 65.6% in combination with contrasting upper body postures and contextual object cues like dirty underpants. Such low recognition rate does in fact provide a large margin for external influences on the face. Indeed, their results show that disgust faces are no longer viewed as

expressing disgust when perceived with an incongruent body. This result is consistent with what has been known for a long time that the effect of the secondary information is the biggest where recognition rates of the primary stimulus are poorest (Massaro & Egan, 1996). This doesn't seem that this study provides good evidence that judgments of facial expressions are *entirely* malleable, since the effects it shows are for facial expressions that are rather ambiguous when they are viewed on their own.

Aviezer et al. (2008) rightly remark that a crucial issue is whether the context effects are post-perceptual rather than truly perceptual (B. de Gelder & Bertelson, 2003). Their experiments unfortunately do not allow a conclusion one way or the other. They did not use rapid presentation or masking, the two classical means of exercising strategic control over the perceptual process. In all experiments they used untimed presentation with free exploration of the compound stimulus which allows the viewer to attend to the face and the body and ultimately to choose what information to base the response on, either on an ad hoc basis or also possibly depending on the particular emotion combination. The eye movement data they recorded do not settle the issue of rapid perceptual procedures in the observer. The eye movements effect they report cannot be deemed to reflect an underlying fast or rapid process, as the fixation latencies to enter either the upper or lower face area are on average around 1000 ms. In view of the fact that the latency to make a saccade is around 150-200 ms the reported latencies are very long indeed. Moreover, comparing their saccade latency values with RTs reported in Meeren et al. (2005) shows that on average RTs are about 200ms faster and even more for the congruent conditions. This is remarkable since RTs are by definition a slower measure than saccades (Bannerman, Milders, de Gelder, & Sahraie, 2009). The findings indicate that the long eye gaze latencies reflect gaze fixation under voluntary-attentional control. Participants look at the compound stimulus and as we have shown, rapidly (in EEG time at the P1, which is in the window around 100ms) realizing the oddity of the compound stimulus and then explore and reassess the facial expression intentionally and apply a verbal label.

In fact, it is easy to imagine the opposite situation where the bodily expression completely loses its categorical expression identity in favor of the facial expression. In view of our limited understanding of what the critical components of bodily expressions are, it is currently still difficult to create stimuli where information from body and face is well balanced with

respect to the informational content such that what each contributes can reliably be compared. More importantly, the relative predominance of the face vs. the body when both are present and are equally attended to may very well depend on the specific emotion. This is already suggested by data from eye movement studies indicating that observers' fixation behavior during perception of bodily expressions is also a function of the emotion displayed. During perception of joy the observers tend to fixate on the head region, whereas during anger and fear most attention is devoted to the hands and arms. For sadness the subjects fixate on heads, arms, and hands and the legs almost never attract the subjects' attention. This fixation behavior is emotion-specific and remains stable under different conditions: whether the subjects were asked to recognize the body postures or were just watching; for both incorrectly and correctly recognized emotions; for pictures with different response times; and during the time progression of the experiment (perceptual learning) (Ousov-Fridin, Barliy, Sheckman, de Gelder, Flash, submitted).

One explanation may be provided by comparing the physical characteristic of different facial expressions. Components of different facial expressions may resemble each other, for example, upturned corners of the mouth characterizes both a smile and pain expression. An example of this strategy is provided by the study just discussed. The role of the context would then be to glue the components together in the configuration reflecting the information from the context. But such a view *prima facie* goes against notion that facial expressions are perceived configurally, and that ERP data indicate that they are rapidly processed.

Emotional expressions involve the whole body in action

Bodyless heads are incomplete visual stimuli just as headless bodies are. To us the body to face context effects primarily suggest not that facial expressions are vague, imprecise or noisy, but that there is a very close link between both. An important aspect to consider when trying to explain that bodily postures influence the expression recognized on a face is provided by recent findings of overarching similarity in the perceptual (configural processing) (Reed *et al.*, 2003) and neurofunctional (spatial and temporal overlap as shown in fMRI, EEG and MEG) signature of facial and bodily expressions (Meeren, Hadjikhani, Ahlfors, Hamalainen, & de Gelder, 2008; Stekelenburg & de Gelder, 2004; van de Riet, Grezes, & de

Gelder, 2009). This suggests that faces as well as bodies can rapidly convey the same message and do so in very similar ways. The brain mentally completes the headless body or the bodyless head. This can obviously not be based on missing physical information as would for example be the case when only part of the face was shown or one component was missing. What triggers the brain's filling in may be, in the case of emotional body postures, the adaptive action the person is engaged in.

From a Darwinian evolutionary perspective, emotions are closely related to actions and therefore likely to involve the whole body rather than only the facial expressions. One view is that emotion provoking stimuli trigger affect programs (Darwin, 1872; Frijda, 1986; Panksepp, 1998; Russell & Feldman Barrett, 1999; Tomkins, 1963), which produce an ongoing stream of neurophysiologic change (or change in a person's homeostatic state) and are associated with evolutionary-tuned behaviors for dealing with stimuli of significant value. Along with the orbitofrontal cortex (OFC) and amygdala, the insula and somatosensory cortex are involved in the modulation of emotional reactions involving the body via connections to brain stem structures (Damasio, 1994, 1999; LeDoux, 1996). This function of the insula and somatosensory cortex may contribute to their important role in emotion perception.

Facial and bodily expressions share largely overlapping neurofunctional basis

Do the results just mentioned indicate that activation to facial expressions and to bodily expressions will almost always show complete overlap? As a matter of fact there is hardly any evidence in the literature to answer this question. For this reason we designed an fMRI study to investigate whether the brain shows distinctive activation patterns for perception of faces and bodies. We presented pictures of faces and faceless bodies that either showed a neutral, fearful or happy expression and asked participants to categorize the emotion expressed by the stimulus. To untangle brain activation related to faces and bodies, we compared how the brain responds to both categories (irrespective of emotional expression). Surprisingly, the results showed that the middle part of the fusiform gyrus (FG) that is typically associated with the perception of facial identity, is more activated for bodies than for faces (van de Riet *et al.*, 2009). Previous studies have shown that there was at least partial overlap between the face-selective and body selective region within the FG (Hadjikhani & de Gelder, 2003; Peelen & Downing, 2005), and van de Riet *et al.* (2009) were the first to directly compare face and

body related activation. In fact, perception of whole body expressions elicited a wider network of brain areas compared to faces, including other areas previously associated with perception of facial expressions, like STS. Other brain regions that were more active for bodies than for faces included the middle temporal/middle occipital gyrus (the so called extrastriate body area, EBA (Downing, Jiang, Shuman, & Kanwisher, 2001), the superior occipital gyrus and the parieto-occipital sulcus. When we consider more specifically the emotional information conveyed by the bodies and faces, again we observed a wider activation pattern specific for emotional bodies than for emotional faces. Interestingly, emotional body expressions activate cortical and subcortical motor areas like caudate nucleus, putamen and inferior frontal gyrus (IFG). This motor related activation may reflect the adaptive action component implied in the body expression, which is less pronounced in facial expressions (B. de Gelder, Snyder, Greve, Gerard, & Hadjikhani, 2004).

Since we used static images in this study, one may argue that the activity in areas associated with movement is related to the fact that there is more implied motion in the body expressions, compared to facial expressions. We therefore performed a follow up study in which we presented video clips of dynamic facial and bodily expressions that conveyed a neutral, fearful or angry expression instead of static picture stimuli. The results were nonetheless in the same line: bodies compared to faces activated more areas than vice versa, including the FG. Again, motor related areas were more activated by emotional body expressions (Kret et al. submitted).

Taken together these findings support the conclusion that while separating perception of faces and bodies may be somewhat artificial, bodily expressions activate a wider network of brain areas, including motor and action related regions.

Facial and bodily expressions in the context of scenes

When observing a stimulus that consists of a face and body with congruent expression (for example a fearful face on a fearful body) one might expect that recognition will be 100% correct. But this is not necessarily the case. In fact, perception and recognition of an emotional action is also influenced by the particular setting or scene in which it occurs. For example, viewed in isolation the sprint to the finish of a man shaking off a competitor looks

quite similar to the flight of a man running away from a robber holding a knife. Without the context information, the emotional valence is ambiguous. Faces and bodies routinely appear as part of natural scenes and our perceptual system seems to be wired to make the link between the expression and the environment in which it appears to us. But little is known about the mechanism underlying this. Older appraisal theories of emotion (e.g. Scherer, Shorr, & Johnstone, 2001) acknowledge the importance of a visual event for our interpretation and evaluation of it and propose explanations for how we (emotionally) react to it. However, the primary focus in appraisal theories regards the emotional response of the observer to a stimulus, rather than the mere perception of the stimulus.

Hierarchical perception models tend to investigate the possible effects of a scene context as semantic effects which occur relatively late and take place in relatively middle to higher cognitive levels of processing (Bar, 2004). However, there is evidence that supports an early perceptual and neuro-anatomical analysis of a scene. Tachistosopic presentation of a scene contributes to subsequent processing of the spatial relations across the scene (Sanocki, 2003), and the rapid extraction of the gist of a scene may be based on low spatial frequency coding (Oliva & Schyns, 1997). The more semantic effects of scene processing occur at a later stage, around 400ms after stimulus onset. For example, objects presented in their usual context are identified better (Davenport & Potter, 2004) and faster (Ganis & Kutas, 2003) and EEG data show the interaction occurs at about 390 ms after stimulus-onset. The functional neuro-anatomy of contextual associations of objects comprises a network including parahippocampal cortex (PHC), retrosplenial cortex, and superior orbital sulcus (Bar, 2004).

However, the effects of the emotional gist of a scene may occur at an earlier level, in line with the evolutionary significance of the information. Few experimental studies currently exist on the influence of emotional scenes on the perception of faces and bodies. In the first explorations of this issue, we presented fearful, disgusted and happy faces embedded in a natural scene (see Figure 3 for an example). The affective valence of the scene was either fearful, disgusted or happy, and the face-scene combinations were emotionally congruent (e.g. fearful face in fearful scene) or incongruent (e.g. fearful face in happy scene). Participants were required to categorize the emotion expressed by the face. The results revealed faster response times and higher accuracies for the congruent stimulus pairs, showing that the emotional expression of a face is recognized better when it is embedded in a congruent scene

(Righart & de Gelder, 2008b). The context effects hold up under different attentional conditions: it can be observed when participants are explicitly decoding the emotional expression of the face (Righart & de Gelder, 2008a) but also when they are primarily focussed on the orientation of the face (Righart & de Gelder, 2006).

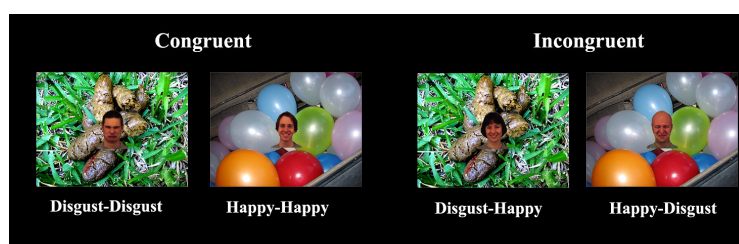


Figure 3. Stimulus examples of congruent and incongruent face-scene compounds (Righart & de Gelder, 2008b)

This indicates that it reflects an early and mandatory process and suggests a perceptual basis. Our EEG studies support this view: when fearful faces are presented in a fearful scene, EEG recordings show a higher N170 amplitude compared to when the same faces are presented in a neutral scene (Righart & de Gelder, 2006).

To investigate how the emotion conveyed by scenes influences brain activity associated with perception of faces we used fMRI while subjects were shown neutral and fearful faces in both neutral and emotional scenes. We ran a parallel version of the experiment with neutral and fearful bodies instead of faces. The results showed that the activation level in FFA is modulated by the kind of scene in which it is presented. In particular, fearful faces elicit more activity than neutral faces, but more interestingly, fearful faces in threatening scenes trigger more activity than fearful faces in neutral scenes. Also, activity in body areas, like the extrastriate body area (EBA) (Downing *et al.*, 2001) is influenced by the scene in which it is embedded: overall, fearful bodies trigger more activity than neutral bodies, but interestingly, neutral bodies in threatening scenes trigger more activity than in neutral scenes. On the other hand, the presence of a face or a body influences brain activity in areas that are associated with the processing of scenes, like the retrosplenial complex (RSC) and the parahippocampal cortex (PHC) (Sinke & de Gelder, submitted; Van den Stock & de Gelder, submitted). In general, neutral scenes trigger higher activation in the PHC and RSC, but the presence of a neutral body boosts activity in these areas. In a behavioural experiment we presented

participants with stimuli depicting an emotional body seen in the foreground against an emotionally congruent or incongruent background. Participants were instructed to categorize the emotion expressed by the foreground body and the results showed that especially negative emotions (fear and anger) were recognized faster in a congruent background, whereas this was not the case for happy expressions (Kret & de Gelder, submitted).

These findings suggest that the emotion conveyed by the scene ‘spills over’ to the embedded face or body, and vice versa. Stated simply, a fearful face makes a neutral scene appear threatening, while a threatening scene makes a neutral face fearful.

Facial expressions in the context of the affective prosody of voices

Research focussing on human face and emotion perception has primarily targeted how visual stimuli are perceived, although in daily life facial expressions are typically accompanied by vocal expressions.

Human emotion recognition can be based on isolated facial or vocal cues (Banse & Scherer, 1996; Scherer, Banse, Wallbott, & Goldbeck, 1991) but combining both modalities results in a performance increase as shown by both increased accuracy rates and shorter response latencies (B. de Gelder, Bocker, Tuomainen, Hensen, & Vroomen, 1999; B. de Gelder & Vroomen, 2000; B. de Gelder, Vroomen, & Teunisse, 1995; Dolan, Morris, & de Gelder, 2001; Massaro & Egan, 1996). Detailed behavioural investigations into crossmodal influences between vocal and facial cues requires a paradigm in which both modalities are combined to create audiovisual pairs. The manipulation ideally consists of altering both the emotional congruency between the two modalities and a task that consists of emotion categorization based on only one of both information streams. For example, de Gelder and Vroomen (2000) presented facial expressions that were morphed on a continuum between happy and sad while at the same time a short spoken sentence was presented. This sentence had a neutral semantic meaning, but was spoken in either a happy or sad emotional tone of voice. Participants were instructed to attend to and categorize the face and ignore the voice in a 2 alternative forced choice task. The results showed a clear influence of the task irrelevant auditory modality on the target visual modality. For example, sad faces were less frequently categorized as sad when they were accompanied by a happy voice. In a follow up experiment, vocal expressions were morphed on a fear-happy continuum and presented with either a fearful or happy face,

while participants were instructed to categorize the vocal expression. Again, the task irrelevant modality (facial expressions) influenced the emotional categorization of the target modality (vocal expressions). Furthermore, this experiment was repeated under different attentional demands, but the facial expression influenced the categorization of vocal expression in every attentional condition (Vroomen, Driver, & de Gelder, 2001).

These findings suggest that affective multisensory integration is a mandatory and automatic process. However, based on these behavioral data, no direct claims can be made about the nature of this crossmodal bias effect. The findings could either reflect an early perceptual or later more cognitive or decisional effect. Neuro-imaging methods with high temporal resolution are needed to provide information on the time course of when this bimodal crosstalk occurs. Studies addressing neural substrates of vocal expressions are few (B. de Gelder, Vroomen, & Pourtois, 2004; George et al., 1996; Ross, 2000) and primarily point to involvement of the right hemisphere. Electroencephalogram (EEG) investigations show that recognition of emotional prosody occurs already within the first 100-150 ms of stimulus presentation (Bostanov & Kotchoubey, 2004; B. de Gelder, Bocker et al., 1999; Goydke, Altenmuller, Moller, & Munte, 2004). The possibility that ecologically relevant audiovisual expressions may rely on specialized neural mechanisms has long been recognized in animal research and several studies have explored the relation between auditory and visual processing streams in non-human primate communication (Ghazanfar & Santos, 2004; Parr, 2004).

EEG studies addressing the time course of audiovisual integration point to an early integration of both modalities (around 110 ms after stimulus presentation) (B. de Gelder, Bocker et al., 1999; Pourtois, de Gelder, Vroomen, Rossion, & Crommelinck, 2000), which is compatible with a perceptual effect. Supporting evidence for a mandatory nature of this integration is provided by studies with blindsight patients, who are unable, due to cortical damage, to consciously perceive visual stimuli presented in a segment of the visual field. When they are presented with auditory vocal expressions and at the same time visual facial expressions in their blind field, fMRI and EEG recordings are influenced by the facial expression of which they are unaware. This shows that the unconscious emotional information displayed by the face is processed by alternative brain pathways through which it influenced the brain responses to the consciously perceived vocal expressions.

Another question concerns where in the brain the integration of perceived vocal and facial expressions takes place. Heteromodal cortex is a logical candidate for multisensory integration (Mesulam, 1998). Superior temporal sulcus (STS) (Barraclough, Xiao, Baker, Oram, & Perrett, 2005) and ventral premotor cortex (Kohler *et al.*, 2002) have been shown to be involved in multisensory integration of biological stimuli. Functional imaging studies addressing the combined perception of emotional face-voice pairs (Dolan *et al.*, 2001; Ethofer *et al.*, 2006) show that fearful faces simultaneously presented with fearful voices activate the left amygdala. The role of the amygdala in emotional and face processing is well established (Zald, 2003) and connectivity data show that it receives inputs from both auditory and visual cortices (McDonald, 1998). These findings make this brain structure an important location for integration of affective bimodal inputs.

Recent studies have shown that next to facial expressions, bodily expressions are also prone to crossmodal affective influences. For example, recognition of dynamic whole body expressions of emotion are influenced not only by both human and animal vocalizations (Van den Stock, Grezes, & de Gelder, 2008), but also by instrumental music (Van den Stock, Peretz, Grèzes, & de Gelder, 2009), suggesting the brain is well organized to combine affective information from different sensory channels.

Summary and conclusions

Real faces are part and parcel of their context and this consideration must play an important role in future models of face processing. Recent data show that bodily expressions, affective prosody, as well as the emotional gist of a natural scene all influence the recognition of facial expression. When a face is accompanied by a body or voice expressing the same emotion, or when it is presented in a congruent emotional scene, the recognition of facial expression typically improves, i.e. both the judgment accuracy and speed increase. Hence, both the immediate visual and auditory contexts function to disambiguate the signals of facial expression. Our behavioral and electrophysiological data suggest that this perceptual integration of information does not require high-level semantic analysis occurring relatively late at higher cognitive centers. Instead, the integration appears to be an automatic and mandatory process, which takes place very early in the processing stream, before full

structural encoding of the stimulus and conscious awareness of the emotional expression are fully elaborated.

References

- Adolphs, R. (2002). Neural systems for recognizing emotion. *Curr Opin Neurobiol*, 12(2), 169-177.
- Adolphs, R., Damasio, H., Tranel, D., Cooper, G., & Damasio, A. R. (2000). A role for somatosensory cortices in the visual recognition of emotion as revealed by three-dimensional lesion mapping. *J Neurosci*, 20(7), 2683-2690.
- Amaral, D. G., & Price, J. L. (1984). Amygdalo-cortical projections in the monkey (*Macaca fascicularis*). *J Comp Neurol*, 230(4), 465-496.
- Andino, S. L., Menendez, R. G., Khateb, A., Landis, T., & Pegna, A. J. (2009). Electrophysiological correlates of affective blindsight. *Neuroimage*, 44(2), 581-589.
- Aviezer, H., Hassin, R. R., Ryan, J., Grady, C., Susskind, J., Anderson, A., et al. (2008). Angry, disgusted, or afraid? Studies on the malleability of emotion perception. *Psychol Sci*, 19(7), 724-732.
- Bannerman, R. L., Milders, M., de Gelder, B., & Sahraie, A. (2009). Orienting to threat: faster localization of fearful facial expressions and body postures revealed by saccadic eye movements. *Proceedings Biological sciences / The Royal Society*, 276(1662), 1635-1641.
- Banse, R., & Scherer, K. R. (1996). Acoustic profiles in vocal emotion expression. *J Pers Soc Psychol*, 70(3), 614-636.
- Bar, M. (2004). Visual objects in context. *Nat Rev Neurosci*, 5(8), 617-629.
- Barracough, N. E., Xiao, D., Baker, C. I., Oram, M. W., & Perrett, D. I. (2005). Integration of visual and auditory information by superior temporal sulcus neurons responsive to the sight of actions. *J Cogn Neurosci*, 17(3), 377-391.
- Bodamer, J. (1947). Die prosop-Agnosie. *Archiv fur Psychiatrie und Nervenkrankheiten*, 179, 6-53.
- Bostanov, V., & Kotchoubey, B. (2004). Recognition of affective prosody: continuous wavelet measures of event-related brain potentials to emotional exclamations. *Psychophysiology*, 41(2), 259-268.
- Carmichael, S. T., & Price, J. L. (1995). Limbic connections of the orbital and medial prefrontal cortex in macaque monkeys. *J Comp Neurol*, 363(4), 615-641.
- Catani, M., Jones, D. K., Donato, R., & Ffytche, D. H. (2003). Occipito-temporal connections in the human brain. *Brain*, 126(Pt 9), 2093-2107.

- Cowey, A. (2004). The 30th Sir Frederick Bartlett lecture. Fact, artefact, and myth about blindsight. *Q J Exp Psychol A*, 57(4), 577-609.
- Damasio, A. R. (1994). *Descartes' Error: Emotion, Reason, and the Human Brain*. New York: Grosset/Putnam.
- Damasio, A. R. (1999). *The Feeling of What Happens*. New York: Harcourt Brace.
- Darwin, C. (1872). *The expression of the emotions in man and animals*. London: John Murray.
- Davenport, J. L., & Potter, M. C. (2004). Scene consistency in object and background perception. *Psychol Sci*, 15(8), 559-564.
- de Gelder, B. (2006). Towards the neurobiology of emotional body language. *Nature Reviews Neuroscience*, 7(3), 242-249.
- de Gelder, B., & Bertelson, P. (2003). Multisensory integration, perception and ecological validity. *Trends Cogn Sci*, 7(10), 460-467.
- de Gelder, B., Bocker, K. B., Tuomainen, J., Hensen, M., & Vroomen, J. (1999). The combined perception of emotion from voice and face: early interaction revealed by human electric brain responses. *Neurosci Lett*, 260(2), 133-136.
- de Gelder, B., Frissen, I., Barton, J., & Hadjikhani, N. (2003). A modulatory role for facial expressions in prosopagnosia. *Proc Natl Acad Sci U S A*, 100(22), 13105-13110.
- de Gelder, B., Meeren, H. K., Righart, R., Van den Stock, J., van de Riet, W. A. C., & Tamietto, M. (2006). Beyond the face: exploring rapid influences of context on face processing. *Progress in Brain Research*, 155, 37-48.
- de Gelder, B., Pourtois, G., van Raamsdonk, M., Vroomen, J., & Weiskrantz, L. (2001). Unseen stimuli modulate conscious visual experience: evidence from inter-hemispheric summation. *Neuroreport*, 12(2), 385-391.
- de Gelder, B., & Rouw, R. (2000). Configural face processes in acquired and developmental prosopagnosia: evidence for two separate face systems? *Neuroreport*, 11(14), 3145-3150.
- de Gelder, B., Snyder, J., Greve, D., Gerard, G., & Hadjikhani, N. (2004). Fear fosters flight: A mechanism for fear contagion when perceiving emotion expressed by a whole body. *Proc Natl Acad Sci U S A*, 101(47), 16701-16706.
- de Gelder, B., & Tamietto, M. (2007). Affective blindsight. *Scholarpedia*, 2(10), 3555.

- de Gelder, B., Tamietto, M., van Boxtel, G., Goebel, R., Sahraie, A., Van den Stock, J., et al. (2008). Intact navigation skills after bilateral loss of striate cortex. *Current Biology*, 18(24), R1128-R1129.
- de Gelder, B., Van den Stock, J., Meeren, H. K., Sinke, C. B., Kret, M. E., & Tamietto, M. (2009). Standing up for the body. Recent progress in uncovering the networks involved in processing bodies and bodily expressions. *Neuroscience and Biobehavioral Reviews*.
- de Gelder, B., & Vroomen, J. (2000). The perception of emotions by ear and by eye. *Cognition and Emotion*, 14(3), 289-311.
- de Gelder, B., Vroomen, J., & Pourtois, G. (2004). Multisensory perception of emotion, its time course and its neural basis. In G. Calvert, C. Spence & B. E. Stein (Eds.), *Handbook of multisensory processes*. (pp. 581-596). Cambridge, MA: MIT.
- de Gelder, B., Vroomen, J., Pourtois, G., & Weiskrantz, L. (1999). Non-conscious recognition of affect in the absence of striate cortex. *Neuroreport*, 10(18), 3759-3763.
- de Gelder, B., Vroomen, J., & Teunisse, J. P. (1995). Hearing smiles and seeing cries. The bimodal perception of emotion. *Bulletin of the Psychonomic Society*, 29, 309.
- Dolan, R. J., Morris, J. S., & de Gelder, B. (2001). Crossmodal binding of fear in voice and face. *Proc Natl Acad Sci U S A*, 98(17), 10006-10010.
- Downing, P. E., Jiang, Y., Shuman, M., & Kanwisher, N. (2001). A cortical area selective for visual processing of the human body. *Science*, 293(5539), 2470-2473.
- Ethofer, T., Anders, S., Erb, M., Droll, C., Royen, L., Saur, R., et al. (2006). Impact of voice on emotional judgment of faces: an event-related fMRI study. *Hum Brain Mapp*, 27(9), 707-714.
- Fernberger, S. W. (1928). False suggestion and the Piderit model. *American Journal of Psychology*, 40, 562-568.
- Fodor, J. (1983). *The Modularity of Mind*. Cambridge, MA: MIT Press.
- Frijda, N. H. (1986). *The emotions*. Cambridge: Cambridge University Press.
- Ganis, G., & Kutas, M. (2003). An electrophysiological study of scene effects on object identification. *Brain Res Cogn Brain Res*, 16(2), 123-144.
- George, M. S., Parekh, P. I., Rosinsky, N., Ketter, T. A., Kimbrell, T. A., Heilman, K. M., et al. (1996). Understanding emotional prosody activates right hemisphere regions. *Arch Neurol*, 53(7), 665-670.

- Ghazanfar, A. A., & Santos, L. R. (2004). Primate brains in the wild: the sensory bases for social interactions. *Nat Rev Neurosci*, 5(8), 603-616.
- Goydke, K. N., Altenmuller, E., Moller, J., & Munte, T. F. (2004). Changes in emotional tone and instrumental timbre are reflected by the mismatch negativity. *Brain Res Cogn Brain Res*, 21(3), 351-359.
- Hadjikhani, N., & de Gelder, B. (2003). Seeing fearful body expressions activates the fusiform cortex and amygdala. *Curr Biol*, 13(24), 2201-2205.
- Haxby, J. V., Hoffman, E. A., & Gobbini, M. I. (2000). The distributed human neural system for face perception. *Trends Cogn Sci*, 4(6), 223-233.
- Haxby, J. V., Horwitz, B., Ungerleider, L. G., Maisog, J. M., Pietrini, P., & Grady, C. L. (1994). The functional organization of human extrastriate cortex: a PET-rCBF study of selective attention to faces and locations. *J Neurosci*, 14(11 Pt 1), 6336-6353.
- Haxby, J. V., Ungerleider, L. G., Horwitz, B., Maisog, J. M., Rapoport, S. I., & Grady, C. L. (1996). Face encoding and recognition in the human brain. *Proc Natl Acad Sci U S A*, 93(2), 922-927.
- Hoffman, E. A., & Haxby, J. V. (2000). Distinct representations of eye gaze and identity in the distributed human neural system for face perception. *Nat Neurosci*, 3(1), 80-84.
- Iidaka, T., Omori, M., Murata, T., Kosaka, H., Yonekura, Y., Okada, T., et al. (2001). Neural interaction of the amygdala with the prefrontal and temporal cortices in the processing of facial expressions as revealed by fMRI. *J Cogn Neurosci*, 13(8), 1035-1047.
- Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: a module in human extrastriate cortex specialized for face perception. *J Neurosci*, 17(11), 4302-4311.
- Kohler, E., Keysers, C., Umiltà, M. A., Fogassi, L., Gallese, V., & Rizzolatti, G. (2002). Hearing sounds, understanding actions: action representation in mirror neurons. *Science*, 297(5582), 846-848.
- LeDoux, J. E. (1996). *The emotional brain: The mysterious underpinnings of emotional life*. New York, NY, US: Simon and Schuster.
- Massaro, D. W., & Egan, P. B. (1996). Perceiving Affect from the Voice and the Face. *Psychonomic Bulletin and Review*, 3, 215-221.
- McDonald, A. J. (1998). Cortical pathways to the mammalian amygdala. *Prog Neurobiol*, 55(3), 257-332.

- McKone, E., & Yovel, G. (2009). Why does picture-plane inversion sometimes dissociate perception of features and spacing in faces, and sometimes not? Towards a new theory of holistic processing. *Psychon Bull Rev*, in press.
- Meeren, H. K., Hadjikhani, N., Ahlfors, S. P., Hamalainen, M. S., & de Gelder, B. (2008). Early category-specific cortical activation revealed by visual stimulus inversion. *PLoS ONE*, 3(10), e3503.
- Meeren, H. K., van Heijnsbergen, C. C., & de Gelder, B. (2005). Rapid perceptual integration of facial expression and emotional body language. *Proc Natl Acad Sci U S A*, 102(45), 16518-16523.
- Mesulam, M. M. (1998). From sensation to cognition. *Brain*, 121 (Pt 6), 1013-1052.
- Morris, J. S., de Gelder, B., Weiskrantz, L., & Dolan, R. J. (2001). Differential extrageniculostriate and amygdala responses to presentation of emotional faces in a cortically blind field. *Brain*, 124(Pt 6), 1241-1252.
- Morris, J. S., Friston, K. J., Buchel, C., Frith, C. D., Young, A. W., Calder, A. J., et al. (1998). A neuromodulatory role for the human amygdala in processing emotional facial expressions. *Brain*, 121 (Pt 1), 47-57.
- Morris, J. S., Ohman, A., & Dolan, R. J. (1998). Conscious and unconscious emotional learning in the human amygdala. *Nature*, 393(6684), 467-470.
- Oliva, A., & Schyns, P. G. (1997). Coarse blobs or fine edges? Evidence that information diagnosticity changes the perception of complex visual stimuli. *Cognit Psychol*, 34(1), 72-107.
- Panksepp, J. (1998). *Affective neuroscience: The foundation of human and animal emotions*. New York: Oxford University Press.
- Parr, L. A. (2004). Perceptual biases for multimodal cues in chimpanzee (*Pan troglodytes*) affect recognition. *Anim Cogn*, 7(3), 171-178.
- Peelen, M. V., & Downing, P. E. (2005). Selectivity for the human body in the fusiform gyrus. *J Neurophysiol*, 93(1), 603-608.
- Peelen, M. V., & Downing, P. E. (2007). The neural basis of visual body perception. *Nat Rev Neurosci*, 8(8), 636-648.
- Peelen, M. V., Lucas, N., Mayer, E., & Vuilleumier, P. (2009). Emotional attention in acquired prosopagnosia. *Soc Cogn Affect Neurosci*, 4(3), 268-277.
- Pegna, A. J., Khateb, A., Lazeyras, F., & Seghier, M. L. (2005). Discriminating emotional faces without primary visual cortices involves the right amygdala. *Nat Neurosci*, 8(1), 24-25.

- Pessoa, L., McKenna, M., Gutierrez, E., & Ungerleider, L. G. (2002). Neural processing of emotional faces requires attention. *Proc Natl Acad Sci U S A*, 99(17), 11458-11463.
- Pourtois, G., de Gelder, B., Vroomen, J., Rossion, B., & Crommelinck, M. (2000). The time-course of intermodal binding between seeing and hearing affective information. *Neuroreport*, 11(6), 1329-1333.
- Puce, A., Allison, T., Asgari, M., Gore, J. C., & McCarthy, G. (1996). Differential sensitivity of human visual cortex to faces, letterstrings, and textures: a functional magnetic resonance imaging study. *J Neurosci*, 16(16), 5205-5215.
- Reed, C. L., Stone, V. E., Bozova, S., & Tanaka, J. (2003). The body-inversion effect. *Psychol Sci*, 14(4), 302-308.
- Righart, R., & de Gelder, B. (2006). Context influences early perceptual analysis of faces--an electrophysiological study. *Cereb Cortex*, 16(9), 1249-1257.
- Righart, R., & de Gelder, B. (2008a). Rapid influence of emotional scenes on encoding of facial expressions: an ERP study. *Social cognitive and affective neuroscience*, 3(3), 270-278.
- Righart, R., & de Gelder, B. (2008b). Recognition of facial expressions is influenced by emotional scene gist. *Cognitive, affective & behavioral neuroscience*, 8(3), 264-272.
- Ross, E. D. (2000). Affective prosody and the aprosodias. In Mesulam, M. Marsel (Ed). (2000). *Principles of behavioral and cognitive neurology* (2nd ed.). (pp. 316 331). London, Oxford University Press. xviii, 540 pp.SEE BOOK.
- Russell, J. A., & Feldman Barrett, L. (1999). Core affect, prototypical emotional episodes, and other things called emotion: dissecting the elephant. *Journal of Personality and Social Psychology*, 76, 805-819.
- Sanocki, T. (2003). Representation and perception of scenic layout. *Cognit Psychol*, 47(1), 43-86.
- Scherer, K. R., Banse, R., Wallbott, H. G., & Goldbeck, T. (1991). Vocal Cues in Emotion Encoding and Decoding. *Motivation and Emotion*, 15(2), 123-148.
- Scherer, K. R., Shorr, A., & Johnstone, T. (2001). *Appraisal processes in emotion: theory, methods, research*. Canary, NC: Oxford University Press.
- Stekelenburg, J. J., & de Gelder, B. (2004). The neural correlates of perceiving human bodies: an ERP study on the body-inversion effect. *Neuroreport*, 15(5), 777-780.

- Tamietto, M., Geminiani, G., Genero, R., & de Gelder, B. (2007). Seeing fearful body language overcomes attentional deficits in patients with neglect. *Journal of Cognitive Neuroscience*, 19(3), 445-454.
- Tomkins, S. S. (1963). Affect, imagery consciousness: Vol. 2. The negative affects. In. New York: Springer verlag.
- van de Riet, W. A., Grezes, J., & de Gelder, B. (2009). Specific and common brain regions involved in the perception of faces and bodies and the representation of their emotional expressions. *Social Neuroscience*, 4(2), 101-120.
- Van den Stock, J., Grezes, J., & de Gelder, B. (2008). Human and animal sounds influence recognition of body language. *Brain Research*, 1242, 185-190.
- Van den Stock, J., Peretz, I., Grèzes, J., & de Gelder, B. (2009). Instrumental music influences recognition of emotional body language. *Brain Topography*, 21(3-4), 216-220.
- Van den Stock, J., Righart, R., & de Gelder, B. (2007). Body expressions influence recognition of emotions in the face and voice. *Emotion*, 7(3), 487-494.
- Van den Stock, J., van de Riet, W. A., Righart, R., & de Gelder, B. (2008). Neural correlates of perceiving emotional faces and bodies in developmental prosopagnosia: an event-related fMRI-study. *PLoS ONE*, 3(9), e3195.
- Vroomen, J., Driver, J., & de Gelder, B. (2001). Is cross-modal integration of emotional expressions independent of attentional resources? *Cogn Affect Behav Neurosci*, 1(4), 382-387.
- Vuilleumier, P., Richardson, M. P., Armony, J. L., Driver, J., & Dolan, R. J. (2004). Distant influences of amygdala lesion on visual cortical activation during emotional face processing. *Nat Neurosci*, 7(11), 1271-1278.
- Vuilleumier, P., & Schwartz, S. (2001). Emotional facial expressions capture attention. *Neurology*, 56, 153-158.
- Zald, D. H. (2003). The human amygdala and the emotional evaluation of sensory stimuli. *Brain Res Brain Res Rev*, 41(1), 88-123.

Chapter 2: Standing up for the body. Recent progress in uncovering the networks involved in processing bodies and bodily expressions²

Introduction

Perception of bodies and bodily expressions is a relatively novel topic in affective neuroscience, a field dominated so far by investigations of facial expressions. Faces and bodies are equally salient and familiar in daily life and often convey some of the same information, leading one to expect that many of the same issues arise about both. Yet research on faces still tends to be dominated by the issue of category specificity. Historically, patients with pure face deficits have proven hard to find and as a consequence the debate on the existence of a pure deficit is still going on. But since the early days of functional brain imaging considerable efforts have been devoted to argue for category specificity as opposed to distributed object representation. Should body researchers follow this well trodden path and focus on uncovering THE body specific brain area(s)? Given the massive evidence positive as well as negative, about categorical representation of faces and the little information we have so far about bodies, it is likely that the case for or against body specificity will remain stronger for faces and this for a long time to come. This is indeed the conclusion reached in a recent paper (Minnebusch & Daum, 2009). Is this the right conclusion?

The goal here is to consider all the available evidence in more detail than done so far but also to argue for a theoretically motivated comparison of faces and bodies that is sensitive to broader issues than only category specificity. Faces, bodies but also many other objects are multidimensional information bearers. Indeed, without a theoretical motivation and a focus on one or a few specific dimensions, it is hard to decide upon the appropriate control category and the results of such somewhat arbitrary comparisons will fail to convince researchers using a different one. For example, emotion expression and action representation are two very important aspects in body representation. The first has an equivalent for faces, but the second

² This chapter is based on de Gelder, B., Van den Stock, J., Meeren, H.K.M., Sinke, C.B.A., Kret, M.E., Tamietto, M. (2010). Standing up for the body. Recent progress in uncovering the networks involved in processing bodies and bodily expressions. *Neuroscience and Biobehavioral Reviews* 34, 513-527.

does not. But even the equivalence concerning the emotional dimension of face and body stimuli may be less straightforward than it seems. We typically relate a facial expression to mental states and feelings, but we tend to associate a bodily expression with the action a person is involved in. So even if from the vantage point of emotion expression they superficially convey the same affective signals, facial and bodily expressions may trigger very different processes in the observer. The facial expression may predominantly produce empathy in the observer while a bodily expression may be a call for action (de Gelder, 2009). And the stimulus equivalent of a body engaged in a neutral action is also difficult to find in the domain of facial information. For example, neutral facial movements often used as control condition are speech, chewing or making a grimace but each triggers other processes than does the observation of a body crossing the street, gardening or playing tennis, all of which count as neutral events but none of which has a clear equivalent in the face domain.

Similarities in perceptual routines: configuration and feature based perception of faces, bodies and other visual objects

The major concepts used to argue for the specificity of perceptual processes involved in face perception are those of configuration vs. feature based processes. Note that configural and holistic are often used interchangeable now, but for a while the notion of holistic processes referred to the strong claim that in some cases of complex stimuli the composing parts or (misleadingly called) the features are not encoded separately (Farah, 1990) and therefore not available for perceptual processes. The more general notion of configural processes does not make the claim that facial features or parts are not represented. So at present there is no reason to refer to holistic processes as different from configural ones (see also below) and it is not clear whether this term will be of any future use in body research. In any case, it is also informative to put the questions about configural processing in the broader context of major theories of object perception and relate them to a classical effect in the visual perception literature, the object superiority effects, generally defined as the influence of whole object recognition on recognition of the parts. Configural processing of faces then appears as a kind of object superiority effect. This puts the debate on configural processing in a broader context because the argument cannot be that faces are special because they are processed configurally. It is worth keeping in mind that the very first reports noting this effect (Yin, 1969) already

used a large number of visual objects besides faces and reported inversion effects for some of them (e.g. landscapes).

To come back to the question raised earlier, similarities between configural processing of faces and bodies were reported as soon as this issue was addressed. There is clear evidence that both faces and bodies are processed configurally, rather than as an assemblage of features. This is typically assessed by measuring the perceptual processes that are triggered by tasks in which faces are presented upside-down (the inversion effect): faces and bodies presented upside-down are relatively more difficult to recognize than inverted objects, like for example houses (Reed, Stone, Bozova, & Tanaka, 2003). By and large evidence from brain imaging studies has since confirmed these first findings of a body inversion effect while also providing information about its time course. Here however, the evidence points to interesting differences with the face inversion effect, as far as time course but also as far as neural generators are concerned (Meeren, Hadjikhani, Ahlfors, Hamalainen, & de Gelder, 2008).

The neurophysiology of body perception. Implications for neural representation and time course of processing

Neurophysiological studies of body processing started with single cell recording studies in the monkey. In the late 1960's Gross et al. (1969; 1972) showed neurons in the inferior temporal cortex (IT) that selectively fired to silhouettes of monkey hands, a result which was later confirmed for drawings of human hands and faces (Desimone, Albright, Gross, & Bruce, 1984). Subsequently, neurons responding to static images of whole bodies, body postures that imply some form of body motion, body orientations, and body movements were discovered in superior temporal sulcus (STS) (Barraclough, Xiao, Oram, & Perrett, 2006; Jellema & Perrett, 2003a, 2003b; Oram & Perrett, 1996; Perrett et al., 1985; Wachsmuth, Oram, & Perrett, 1994).

Strong evidence for body shape specificity was found recently by Kiani et al. (2007) who recorded from a large population of single cells in monkey IT. The response pattern of the cell population as a whole was far more category-selective than that of single cells, implying that not only cell responses to the preferred category but also responses to the suboptimal

categories carry important information. The population responses formed category clusters that resembled our intuitive object categories, with a first main division between animals and inanimate objects. The animal-characteristic responses could be divided into face- and body-related responses, and these could be further subdivided on the basis of certain animal classes, with the body population code forming three distinct clusters, i.e. for human bodies, four-limbed animals and birds, and lower animals.

Intracranial recordings from the ventral and lateral temporal cortex in epileptic patients showed a first hint of body-related processing in the human brain with distinct electrode sites being responsive only to photographs of hands and not faces or other objects (McCarthy, Puce, Belger, & Allison, 1999). Since there is strong evidence for distinct representations of faces, bodies and hands with whole bodies and hands forming completely non-overlapping clusters (Kiani et al., 2007), we will further focus on whole body-selective processing and will not discuss studies on hand perception. Using a subdural grid on the lateral occipital cortex, Pourtois et al. (2007) observed highly spatially-specific body-selective responses in the human brain. One electrode site positioned at the junction of the middle temporal and middle occipital gyrus (see table 1 and 2 to compare coordinates) showed strong local field potentials peaking at 260 ms elicited by photographs of whole human bodies (without heads) and not by faces, animals or tools.

Further information on the time-course of body-selective processing in the human brain has been obtained from non-invasive electrophysiological recordings. The deflections in the Event Related Potentials (ERP) of face and body perception show several similarities (Gliga & Dehaene-Lambertz, 2005; Meeren, van Heijnsbergen, & de Gelder, 2005; Righart & de Gelder, 2007; Stekelenburg & de Gelder, 2004; Thierry et al., 2006). ERPs for faces as well as for bodies show a P1 and a prominent N1 component with similar scalp topography (Stekelenburg & de Gelder, 2004). The N1, better known as the “N170” in the case of face processing, a negative deflection at occipitotemporal electrodes peaking between 140-220 ms post stimulus onset, is thought to reflect a late stage in the structural encoding of the visual stimulus (Bentin, Allison, Puce, Perez, & et al., 1996; Eimer, 2000). The mean peak latency of the N1 component for body processing has been found to range between 154 and 228 ms after stimulus onset (Gliga & Dehaene-Lambertz, 2005; Meeren et al., 2005; Minnebusch & Daum, 2009; Righart & de Gelder, 2007; Stekelenburg & de Gelder, 2004; Thierry et al.,

2006; van Heijnsbergen, Meeren, Grezes, & de Gelder, 2007), similar as found for faces (see Table 1).

When faces and bodies are directly compared, the peak latency of the N1 for whole human bodies that include heads (with faces masked) was found to be faster than that for faces (Meeren et al., 2005; Righart & de Gelder, 2007; Stekelenburg & de Gelder, 2004). When headless bodies are presented however, the N1 response slows down to become slower than that for faces (Gliga & Dehaene-Lambertz, 2005; Thierry et al., 2006). When analyzed at a higher spatial resolution, the body and face N1 showed a slightly different spatial pattern, both in their potential distribution on the scalp (Gliga & Dehaene-Lambertz, 2005) and their corresponding source localizations in the brain (Thierry et al., 2006). For both stimulus categories, the sources were located in the right posterior extrastriate cortex, but with a slightly more dorsal distribution for bodies as compared to faces. Different underlying neural generators for face and body perception in the N1 time-window were recently confirmed by us using magnetoencephalography (MEG) with anatomically-constrained distributed source modeling (Meeren et al., submitted). The ventral inferior temporal cortex, including middle fusiform gyrus (FG) showed strong differential activation to face stimuli compared to bodies and different classes of control stimuli (i.e. scrambles and houses) between 120-180 ms after picture onset. Body stimuli, on the other hand, elicited selective responses in the right lateral occipitotemporal cortex, a location corresponding to the EBA. No evidence could be found for early activation of the ventral temporal cortex during body perception. These neuromagnetic findings strongly argue against the proposed functional analogies between the face-sensitive and body-sensitive areas in the FG (Minnebusch & Daum, 2009; Taylor, Wiggett, & Downing, 2007).











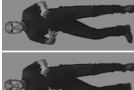


The well-known electrophysiological inversion effect for faces, i.e. an increase in amplitude and latency of the N170 has also been found for bodies (Minnebusch & Daum, 2009; Righart & de Gelder, 2007; Stekelenburg & de Gelder, 2004; Taylor et al., 2007). The earlier inversion effect as observed for faces on the P1 component (~120 ms), could however not be found for bodies (Righart & de Gelder, 2007). Note in this context that the inversion effect needs to be assessed as the relative difference in latency and amplitude between a given stimulus and its upside down presented counterpart. Because of the sensitivity of ERP to physical stimulus differences direct comparisons between faces and bodies are misleading.

Adopting that criterion we see that the inversion effect is of the same magnitude for faces and bodies (Stekelenburg & de Gelder, 2004).

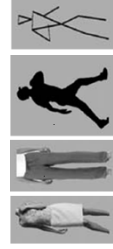
This matter may be pursued by taking advantage of the sensitive time measurements that MEG provides. We recently investigated the earliest onset of the electrophysiological inversion effect for face and body stimulus categories (Meeren et al., 2008). Anatomically-constrained distributed source analyses revealed that both faces and bodies already show inversion effects between 70-100 ms post stimulus with larger responses for the inverted images. Interestingly the cortical distribution of this early inversion effect was highly category-specific. For faces it was found in well-known face-selective areas (e.g. the right inferior occipital gyrus (IOG) and mid FG), whereas for bodies it was found in the posterior-dorsal medial parietal areas (the precuneus / posterior cingulate). Hence, whereas face inversion modulates early activity in face-selective areas in the ventral stream, body inversion evokes differential activity in dorsal stream areas, suggesting different early cortical pathways for configural face and body perception, and again different time courses of activation in the common neural substrate in the FG.

Taking together all currently available information on the time course of body and face processes brings us to the conclusion that reports of time courses and a fortiori comparative ones of different visual objects cannot be confined to the presence/absence of a pre-defined marker (e.g. the face specific N170). We need to look at different time windows in different brain areas, some of which also activate during more than a single window.

Table 1. Overview of the N170 ERP component in body perception studies: stimuli, task, findings. All findings are from normal adults unless stated otherwise. Results on other ERP components (except VPP) and results in infants (Gliga & Dehaene-Lambertz, 2005) are not included here. EEG = electroencephalography; LFP = local field potentials; AMP = Amplitude; LAT = Latency; TOP = Topography; S_LOC = Source Localisation; LOC=Location. RH = right hemisphere; LH = left hemisphere.

Study	Body Stimuli (presentation duration)	Non-body stimuli	Task	N170 findings
Scalp EEG				
Stekelenburg & De Gelder, 2004	<div>   </div> <div>500ms (faces blurred)</div>	<div>    </div>	Forced-choice upright vs inverted judgement	AMP: Faces>Bodies; Bodies=Shoes. LAT: Bodies (157ms) < Faces (167ms) TOP: Bodies and Faces very similar, different for Shoes
Gliga & Dehaene-Lambertz, 2005	<div>   </div> <div>1500ms</div>	<div>    </div>	Passive viewing	Inversion Effect: Increase in LAT (12 ms) and AMP for both Faces and Bodies AMP: Bodies = Faces LAT: Bodies (228ms) > Faces (204 ms) TOP: different for faces and bodies
Meeren <i>et al.</i> , 2005	<div>   </div> <div>200 ms emotionally congruent and incongruent face-body compound stimuli</div>		Forced-choice emotion classification of the face (on the body for headless bodies)	Distorted bodies: Decrease in AMP, no LAT effect AMP: Bodies = Faces LAT: Face-Body compounds (177ms) < Faceless Bodies (187ms) < Faces (199ms) Effects of Bodily Emotion: LAT Angry Bodies (175 ms) < Fearful Bodies (180 ms) in Face-Body Compound stimuli. No effect for Faceless Bodies or Faces.

Thierry *et al.*,
2006



1-back repetition
detection task

AMP: Faces > Bodies > Objects & Scenes
LAT: Bodies (Photos 190ms / Silhouettes 185ms / Stick figures 193ms) > Faces (170ms)

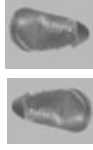
TOP: Different for Faces vs. Bodies
(photos/silhouettes/stick figures)

S_LOC: right posterior extrastriate cortex; larger
and more dorsal for Bodies than for Faces.

200 ms



500 ms
(faces blurred)



Delayed forced-
choice upright vs
inverted
judgement

AMP: Faces > Bodies

LAT: Upright Bodies (155ms) < Upright Faces
(165ms) < Upright Shoes (179 ms)

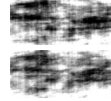
Inversion Effect: Increase LAT and AMP for both
Faces (13ms) and Bodies (9ms).

Opposite inversion effect in prosopagnosics for
Faces (LAT & AMP in 2/4 patients) and for
Bodies (LAT in 2/4, AMP 3/4 patients)

Van
Heijnsbergen
et al., 2007



700 ms
Neutral and Fearful
Body postures
(faces blurred)



Catch trial
detection (white
star
superimposed on
body)

NI70/VPP

AMP: Bodies >> Scrambles (absent)

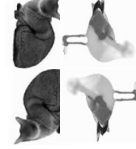
LAT: Bodies: 175ms (NI70) / 177ms (VPP)

Fear effect: Decrease in VPP LAT (4ms)

Minnebusch
et al., 2009







250 ms
Bodies with
and without
heads (faces
blurred)



Same/different
judgement on two
consecutive
stimuli
+ Houses
(no examples
shown)

AMP: Human & Animal Bodies >> Houses.
LAT: No category effect.

Inversion Effect: Increase LAT (6 ms) & AMP for
Bodies with Heads; *Opposite inversion effect*
AMP for Headless Bodies and Cats

Intracranial LFP					
Pourtois <i>et al.</i> , 2007		500 ms	  	Four-alternative forced choice object categorization	<p>Highly spatially-specific body-selective responses on one electrode location within subdural grid.</p> <p>AMP: Headless Bodies >> Faces, Animals, Tools.</p> <p>LAT: surface negative component peaking at 260 ms.</p> <p>LOC: Talairach coordinates responsive electrode (compare with Table 2): x = +39, y = -77, z = -2</p>





Similarities and differences in neurofunctional basis of faces and bodies





This last decade brain imaging has been the method of choice for researchers interested in category specificity. There is evidence that apart from face-selective areas there exist body-selective areas in the brain. In monkeys functional Magnetic Resonance Imaging (fMRI) revealed specialized patches for faces and bodies within the object-selective cortex (Pinsk, DeSimone, Moore, Gross, & Kastner, 2005; Tsao, Freiwald, Knutsen, Mandeville, & Tootell, 2003). In addition, faces and bodies, like the other categories tested, elicited unique distributed response patterns outside these specialized patches (Tsao et al., 2003). Whether these fMRI-defined body-areas consist entirely of body-selective cells as was the case for faces (Tsao, Freiwald, Tootell, & Livingstone, 2006) remains to be investigated. An important comparative study that builds on the single cell recording study of Kiani et al. (2007) provided evidence for the similarities in object representation in the monkey and the human brain (Kriegeskorte et al., 2008). Computations of dissimilarity for distributed blood oxygenation level dependent (BOLD) response patterns in the human brain revealed that these response patterns form category clusters corresponding to animate and inanimate objects. Within the former, subclusters are formed by faces and bodies, and the response patterns distinguish similarly between within-category exemplars in the monkey and the human brain.




Two areas in the body perception network have been the target of categorical selectivity research. The one reported first was an area at the junction of the middle temporal and middle occipital gyrus, labeled the extrastriate body area (EBA) (Downing, Jiang, Shuman, & Kanwisher, 2001). A later added one is in the FG, at least partly overlapping the so-called fusiform face area (FFA) (Kanwisher, McDermott, & Chun, 1997) and termed the fusiform body area (FBA) (Peelen & Downing, 2005b) (see Table 2). Note that as the findings about FBA suggest, face and body category representation may be very closely related. It is a reasonable question whether the sight of a body activates face representations and vice versa because it is well known that when shown familiar stimuli presented in an unusual fashion the brain will automatically fill in the missing information. This issue is important for the notion of category specific representation and has not extensively been addressed (but see Cox, Meyers, & Sinha, 2004).

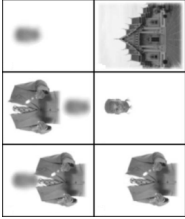


There are only a few fMRI investigations in humans presenting a direct comparison between faces and bodies. In one such study, subjects were scanned while they viewed static images of neutral and emotional (fearful and happy) faces and bodies and categorized the emotion expressed by the stimulus, irrespective of category. When we compared the hemodynamic brain responses of faces vs. bodies (irrespective of emotional expression) we observed that perception of bodies triggers a broad network of brain areas, including areas previously associated with perception of faces, like the FG, STS, and the middle occipital gyrus (MOG) (van de Riet, Grezes, & de Gelder, 2009). Next to these, also the superior occipital gyrus (SOG), the parieto-occipital sulcus (POS) and the intraparietal sulcus responded more to bodies than to faces. Areas more responsive to faces than bodies were restricted to the calcarine sulcus, cerebellum, superior frontal gyrus and anterior cingulate gyrus. Other studies typically focus on the FG and compare areas that respond more to bodies than to objects with areas that respond more to faces than to objects (Peelen & Downing, 2005b; Schwarzlose, Baker, & Kanwisher, 2005; Spiridon, Fischl, & Kanwisher, 2006) rather than a direct comparison between faces and bodies. Using this indirect comparison approach, a previous study by Schwarzlose et al. (2005) with 7 subjects (in 5 of which a right FBA could be defined) found an effect of scanning resolution on the selectivity for either faces or bodies in the FG, with a positive correlation between selectivity and resolution. Their results showed a higher selectivity for faces in the FG, which is compatible with the results of a subsequent study of the same lab (Spiridon et al., 2006), but shows the inverse pattern of our results. This may be related to methodological issues, such as scanning parameters, applied contrasts or the different object categories that were presented.








Other issues hamper a direct comparison of faces and bodies. First, the face is only a part of an object (person), whereas a body (even with the facial area blurred) constitutes a complete object. This may play a role when comparing faces with objects, but also when making the indirect comparison: faces vs. objects compared to bodies vs. objects. Second, perception of faces allows a detailed and fine-grained analysis of the tonus of the complex musculature of the face, whereas bodies are generally covered with clothing, which may conceal important information about muscle tension, especially when dealing with emotional expressions. Also, the presence of clothing implies perception of man-made objects, whereas faces are typically presented in isolation.






Table 2. Overview fMRI studies on whole body perception: stimuli, task, stereotactic coordinates for EBA and FBA and experimental results. ASD = Autism Spectrum Disorder; Ant = anterior; Post = posterior; l_EBA = left EBA; r_EBA = right EBA; l_FBA = left FBA; r_FBA = right FBA. All stimuli are static unless stated otherwise: (1) Dynamic and Static Stimuli, (4) Dynamic stimuli; (2) coordinates based on comparison of observation of dynamic vs static legs; (3) study focused on MT/V5, so no conclusions about EBA are available from the study; (5) coordinates based on comparison of observation of proximal motion vs observation of distal motion; (6) coordinates based on comparison of observation of axial motion vs observation of distal and proximal motion; (7) coordinates based on comparison of observation of neutral hands vs control						
	Localizer focus stimuli	Localizer contrast stimuli	Experimental stimuli	experimental task (localizer task)	Tal coordinates	Main experimental results in EBA/FG
Downing <i>et al.</i> , 2001				(1-back repetition detection)	l_EBA: -50, -69, 11 r_EBA: 50, -69, 4	
Grossman and Blake, 2002			Biological motion, scrambled motion, whole body motion	1-back repetition detection (1-back repetition detection)	l_EBA: -39, -70, 14 r_EBA: 41, -66, 11	EBA not differentially active when observing static headless bodies, dynamic whole bodies (with faces) and faces. EBA not differentially active when observing biological motion compared to scrambled motion.

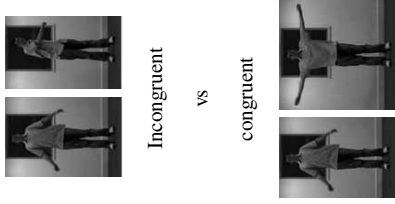
Hadjikhani and de Gelder, 2003			 fearful  instrumental	Passive viewing	L_ FBA: -34, -55, -13 r_ FBA: 35, -55, -14	FFA more activated by observing fearful bodies than instrumental bodies
de Gelder <i>et al.</i> , 2004			 fearful  instrumental	Passive viewing	L_ FBA: -32, -51, -14 r_ FBA: 35, -60, -12	FG more activated by observing fearful bodies than instrumental bodies

Chan <i>et al.</i> , 2004	Bodies	Faces, tools and scenes	 <p>egocentric vs allocentric body parts</p>	<p>1-back repetition detection</p> <p>(1-back repetition detection)</p>	<p>L_ EBA: -45, -73, 4 r_ EBA: 46, -69, 1</p>	<p>Right EBA (not left EBA) more activated by observing allocentric body parts than egocentric body parts. No modulation by self vs other body parts</p>
Astafiev <i>et al.</i> , 2004				<p>Visually guided self generation of finger movement</p> <p>(1-back repetition detection)</p>	<p>L_ EBA: -45, -65, 12 r_ EBA: 42, -58, 13</p>	<p>EBA activated by self-generated finger movements.</p>




Cox <i>et al.</i> , 2004				1-back repetition detection		FFA only activated by observing bodies with spatially congruent degraded face
Wheaton <i>et al.</i> , 2004			 Faces hands legs ⁽¹⁾	Passive viewing	r_FBA: 40, -57, -20 ⁽²⁾	FG more activated by observing dynamic than static legs (not hands) ⁽³⁾
Sakreida <i>et al.</i> , 2005			 fingers, mouth, knee, ankle, elbow, wrist, trunk, and shoulder ⁽⁴⁾	Motion acceleration monitoring	L_EBA: -37, -73, 14 r_EBA: 41, -70, 9 ⁽⁵⁾ L_EBA: -38, -76, 17 r_EBA: 40, -64, 18 ⁽⁶⁾	EBA more activated when viewing motion of proximal body parts than distal body parts. EBA more activated when viewing motion of axial body parts than proximal and distal body parts.

Peelen and Downing, 2005a		Faces, tools and scenes		Visually guided self generation of finger movement	L_EBA: -45, -74, -3 r_EBA: 48, -68, 0	EBA activated by self-generated finger movements. Body selectivity in EBA not correlated with self generated movement response
Peelen and Downing, 2005b			 Stick bodies vs control	1-back repetition detection (1-back repetition detection)	r_EBA: 45, -65, 2 r_FBA: 40, -43, -17	FBA and FFA show partial overlap. EBA and FBA, but not FFA more activated by observing stick figure bodies than scrambled controls.
Peelen and Downing, 2005c		  		(1-back repetition Detection)	L_EBA: -43, -72, -2 r_EBA: 46, -70, -1	Between and within session functional localization reproducibility estimate for EBA





Schwarzlose <i>et al.</i> , 2005				(1-back repetition Detection)	FBA [No coordinates given]	FBA partly overlaps with FFA. Higher body selectivity in FBA at higher spatial scanning resolution. Non-overlapping FBA voxels show no preference for faces.
de Gelder and Hadjikhani, 2006			 Happy vs instrumental			EBA and FBA not differentially activated for observing happy and instrumental bodies presented in the blind field of a hemianope patient.
Downing <i>et al.</i> , 2006a		Heterogeneous set of objects (only few presented here) 		Passive Viewing	L_ EBA: -45, -69, 4 r_ EBA: 45, -67, 5	r_ EBA more activated by bodies than any other category l_ EBA shows similar activation for bodies and birds

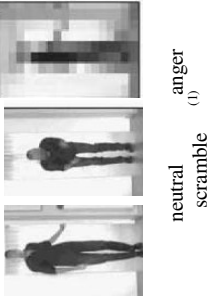
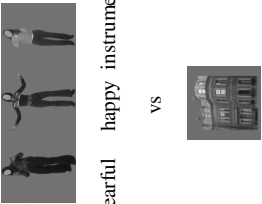
Downing <i>et al.</i> , 2006b	Headless bodies	Faces, tools and scenes	 <p>Incongruent vs congruent</p>	passive viewing (1-back repetition detection)	<p>r_ EBA: 48, -70, 0</p> <p>r_ FBA: 41, -43, -19</p>	EBA more activated by observing video frames of different actions than when observing successive video frames of a single action
Peelen <i>et al.</i> , 2006	Headless bodies	For EBA: Faces, tools and scenes For FBA: Tools	<p>Biological motion vs scrambled motion</p>	passive viewing (1-back repetition detection)	<p>l_ EBA: -45, -74, -1</p> <p>r_ EBA: 48, -70, 1</p> <p>r_ FBA: 41, -45, -19</p>	EBA and FBA sensitivity to bodies correlates with sensitivity to biological motion




Morris <i>et al.</i> , 2006			 background body only body-face	 face only	Passive viewing	L_EBA: -42, -80, 12 r_EBA: 54, -79, 12	EBA less activated when a face is visible. FG more activated when observing bodies with faces than bodies only or faces only.
Saxe <i>et al.</i> , 2006	Headless bodies and body parts	Objects and object parts	 Egocentric vs Allocentric body parts		l-back repetition Detection (passive viewing)	L_EBA: -45, -71, 7 r_EBA: 55, -65, 7	Right EBA (not left EBA) more activated by observing allocentric body parts than egocentric body parts.


Spiridon <i>et al.</i> , 2006	Body parts	Man-made objects or food			Ant. l_EBA: -58, -70, 8 Post. l_EBA: -63, - 80, 24 Ant. r_EBA: 32, -66, -2 Post. r_EBA: 32, -75, 6	FFA not differentially activated when observing faces and body parts
Grosbras and Paus, 2006			   <p>Neutral hands angry hands</p> <p>Neutral face angry face control a⁽⁴⁾</p>	Passive viewing	l_EBA: -51, -64, 5 r_EBA: 48, -63, 5 l_FBA: -41, -53, -12 r_FBA: 43, -49, -12 ⁽⁷⁾	Right FG not differentially activated when observing neutral hands and neutral faces. Left FG more activated when observing neutral hands than neutral faces. EBA activated when observing neutral and angry faces.

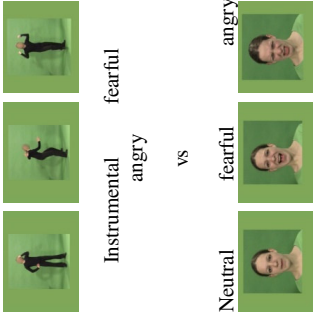
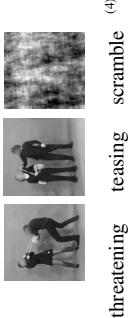
<p>Grezes <i>et al.</i>, 2007</p>			<p>neutral fearful scramble⁽¹⁾</p>	<p>Oddball detection (inverted stimulus)</p>	<p>L_ EBA: -44, -79, 8 r_ EBA: 51, -69, 9</p> <p>L_ FBA: -40, -55, -12 r_ FBA: 44, -51, -16</p>	<p>EBA more activated when observing dynamic than static bodies. Right EBA more activated when observing fearful than neutral bodies.</p> <p>FBA more activated when observing static than dynamic bodies</p>
<p>Peelen <i>et al.</i>, 2007</p>	<p>Headless bodies</p>	<p>Tools</p>	<p>Emotional body expressions⁽¹⁾</p>	<p>emotion intensity rating (1-back repetition detection)</p>	<p>L_ EBA: -47, -71, 3 r_ EBA: 48, -66, 3</p> <p>r_ FBA: 39, -44, -16</p>	<p>EBA and FBA modulated by emotion</p>



Taylor <i>et al.</i> , 2007	Headless bodies & body parts	Chairs & object parts	 Increasing body amount vs Increasing object amount 	1-back repetition detection (1-back repetition detection)	L_EBA: -47, -71, 0 r_EBA: 49, -65, 4 r_FBA: 38, -41, -17	Increasing activation in EBA with increasing body amount. No gradual increase in FBA
Downing <i>et al.</i> , 2007				1-back repetition detection	L_EBA (whole body): -51, -68, 3 r_EBA (whole body): 51, -67, 3 L_EBA (body parts): -48, -67, 6 r_EBA (body parts): 48, -70, 3	

Pichon <i>et al.</i> , 2008			 <p>neutral scramble anger (1)</p>	Oddball detection (inverted stimulus)	<p>L_ EBA: -44, -83, 4 r_ EBA: 50, -74, 4 L_ FBA: -40, -42, -16 r_ FBA: 42, -50, 1</p>	EBA and FBA more activated when observing dynamic than static bodies. EBA and FBA more activated when observing angry than neutral bodies.
Van den Stock <i>et al.</i> , 2008			 <p>fearful happy instrumental vs</p>	Oddball detection (chair)	<p>r_ EBA: 45, -63, 4 r_ FBA: 44, -40, -15</p>	EBA of developmental prosopagnosics more activated when observing neutral faces than control subjects.

Lamm and Decety, 2008	Headless bodies (from Downing, 2001)	Chairs	 <p>intact body part vs pain inflicted body part</p>	Evaluation of amount of pain infliction (1-back repetition detection)	<p>L_ EBA: -53 -65 12 r_ EBA: 53, -65, 11</p>	EBA not differentially activated when observing pain infliction on perceived body parts than unharmed body parts.
van de Riet <i>et al.</i> , 2009			 <p>fearful happy instrumental vs fearful happy neutral</p> 	Emotion categorization	<p>L_ EBA: -50, -72, 7 r_ EBA: 53, -64, 9</p>	<p>EBA is not differentially activated when observing fearful or happy bodies than instrumental bodies.</p> <p>FG more activated by bodies than by faces and more by happy bodies.</p>

Hadjikhani <i>et al.</i> , 2009			 fearful vs instrumental	Passive viewing		FG more activated when observing fearful bodies than instrumental bodies in normals. No differential activation in FG for fearful vs instrumental bodies in ASD
Hodzic <i>et al.</i> , 2009	Headless bodies in dark bathing suit	Neutral objects	Bodies in bathing clothes, including participants own bodies	Recognition of own body (1-back repetition detection)	l_ EBA: -43, -73, -7 r_ EBA: 42, -63, 10 r_ FBA: 38, -46, -13	FBA and left EBA more activated when observing self vs other body parts.
Peelen <i>et al.</i> , 2009	Headless bodies (Peelen and Downing, 2005a)	Scenes and tools (Peelen and Downing, 2005a)	dynamic emotional body expressions (Peelen <i>et al.</i> , 2007)	emotion intensity rating (1-back repetition detection)	l_ EBA: -48, -67, -8 r_ EBA: 39, -76, 7 r_ FBA: 42, -49, -20	Right FBA and EBA in an acquired prosopagnosic show no emotional modulation.

Kret <i>et al.</i> (submitted)			 <p>Instrumental angry</p> <p>vs</p> <p>Neutral fearful</p> <p>angry</p>	Oddball detection	<p>L_ EBA: -50, -72, 7</p> <p>r_ EBA: 50, -72, 7</p>	FG more activated by bodies than by faces. EBA more activated when observing dynamic fear and anger expressions than neutral expressions
Sinke <i>et al.</i> , (in press)			 <p>threatening</p> <p>teasing</p> <p>scramble ⁽⁴⁾</p>	Emotion categorization & color categorization	<p>L_ EBA: -46, -68, 4</p> <p>r_ EBA: 45, -65, 1</p> <p>L_ FBA: -39, -44, -15</p> <p>r_ FBA: 37, -43, -13</p>	EBA and FBA more activated when observing threatening vs teasing human interactions. EBA and FBA more activated when performing interaction categorization than rapid dot detection

Van den Stock <i>et al.</i> , (in prep)				(1-back repetition detection)	<p>L_EBA: -47, -65, 16 r_EBA: 49, -65, 7</p> <p>L_FBA: -39, -42, -19 r_FBA: 40, -41, -12</p>	
---	---	---	--	-------------------------------------	--	--

Emotional modulation of body selective processing

The studies discussed so far used still images as well as static bodies, i.e. bodies not engaged in any activity nor showing an emotional expression. As the case of faces illustrates, there are now many reports showing that facial expressions trigger activations that are earlier and spatially different from those typically obtained previously with neutral faces. These findings are clearly important because they challenge accepted traditional models of face processing which typically start from structural face encoding and associate this with the e.g. the time window of N170.

Using ERP recordings we found evidence for very fast automatic processing of bodily expressions. Images of fear expressions compared to neutral bodies shown as performing the same action already affected the ERP responses in the earliest stage of visual perception, i.e. the P1 component (van Heijnsbergen et al., 2007). In addition, this component showed sensitivity to emotional congruence of the body and face (Meeren et al., 2005). Using MEG and anatomically realistic distributed source modeling we were able to confirm and extend these first results (Meeren et al., in preparation). Within the first 120 ms after picture onset, bodies expressing fear differentially activated the occipital pole, regions in the anterior temporal lobe, the parieto-occipital sulcus, and the intraparietal sulcus as compared to neutral bodies.

Similar to studies that reported emotional modulation of face specific areas, like FFA and OFA (occipital face area) (e.g. Breiter et al., 1996; van de Riet et al., 2009; Vuilleumier, Armony, Driver, & Dolan, 2001), effects of emotional information on activation levels of body areas in the brain have been investigated. The first study addressing this issue focused on the FG and amygdala (AMG) and found an increased activation for fearful bodies compared to instrumental bodies in both areas (Hadjikhani & de Gelder, 2003). This result was very similar to the most frequently observed result for faces and consistent with a connection AMG-FG. In a follow up experiment, we focused on the whole brain and compared the activation of fearful, happy and neutral body expressions (de Gelder, Snyder, Greve, Gerard, & Hadjikhani, 2004). The results confirmed our previous findings, but showed additionally the involvement of motor areas in the perception of emotional body expressions.

Similar results were obtained in a direct comparison of neutral and emotional faces and bodies (van de Riet et al., 2009) in which we also observed that emotional bodies activate cortical and subcortical motor related structures, like inferior frontal gyrus (IFG), caudate nucleus and putamen. These motor areas were not active for emotional faces (van de Riet et al., 2009). Results of emotional modulation of EBA are less clear. We did not observe a difference between neutral and emotional bodies using static body images (van de Riet et al., 2009), but our data with dynamic body expressions show clear emotional modulation of EBA (Grèzes, Pichon, & de Gelder, 2007; Kret, Pichon, Grèzes, & de Gelder, submitted; Pichon, de Gelder, & Grèzes, 2008; Sinke, Sorger, Goebel, & de Gelder, 2009). Recently, we demonstrated in monkeys, that perceiving body postures of conspecifics communicating threat increases the BOLD signal in a subset of body-sensitive voxels in the STS (de Gelder & Partan, 2009).

Static vs. Dynamic body perception

All studies discussed so far used static facial and bodily pictures. But in real life, we are confronted with faces and bodies in motion. Although static body pictures may imply motion, dynamic stimuli contain explicit movement information.

Studies of bodily expressions reported better recognition rates for dynamic than static emotional body stimuli (Atkinson, Dittrich, Gemmell, & Young, 2004; de Meijer, 1989). Two studies designed to find out specifically what additional information is contributed by the dynamics used 3s video clips showing a person opening a door in either a neutral, a fearful (Grèzes et al., 2007) or an angry (Pichon et al., 2008) way, and compared them to still snapshots taken from the same video clips. Not surprisingly, the dynamic body expressions vs. the static expressions (irrespective of the emotional content) triggered activity in motor areas (bilateral activations of the premotor cortex (PM). Emotion processing areas (parietal cortex, STS and FG), but also EBA were active when comparing the emotional vs. the neutral stimuli, irrespective of whether they were presented in a dynamic or static mode. But the most interesting finding was an interaction between emotion and motion in STS and right PM, with stronger motion-related activation for the actions expressing fear as compared to the neutral

actions. Interestingly, STS, PM and parietal cortex were activated during action observation and could be involved in action understanding (Grèzes & Decety, 2001). STS and PM may also be essential for bodily emotion understanding since they seem to represent the emotional action.

A recent fMRI study takes research on body perception one step further by investigating the neurofunctional basis of observing interactions between two people. Sinke et al. (in press) used 3s video clips in which a male actor tries to grab the handbag of a female in either a threatening way (by which the girl is really frightened) or in a teasing way (as if they know each other). During each movie, three small dots appear each for 40ms. In one task condition, participants had to categorize the interaction as threatening or teasing, and in the other they had to ignore the bodies while monitoring the randomly appearing dots and categorize their color, this to try to get their attention away from what really goes on in each situation. This study showed that the right AMG reacts in both cases more to the threatening than teasing movies. When the observer's attention is not explicitly directed to the interaction, this goes together with heightened activation in body sensitive visual regions in FG, posterior middle temporal gyrus and STS. In line with this activation pattern, participants showed a better behavioral performance during the threatening unattended interactions. Furthermore, regions involved in action observation (IFG, temporo-parietal junction (TPJ) and inferior parietal lobe) and preparation (PM, putamen) showed activation for threatening videos, even though the threat is not directed towards the observer.

Following up on the direct comparison of still pictures of faces and bodies by van de Riet et al. (2009) we used dynamic stimuli (Kret et al., submitted) to record female and male participants' hemodynamic brain activity while they observed short video fragments in which female and male actors expressed either fear, anger or neutral signals with the face or body. The AMG was sensitive to all expressions, but significantly more to faces. The areas for processing bodies included FG, EBA and STS. Besides that, we observed activation for threatening body expressions, whether fearful or angry, in EBA, right TPJ and the somatosensory cortex. Whereas studies using static stimuli failed to find evidence for emotion modulation in EBA (de Gelder et al., 2004; van de Riet et al., 2009), dynamic emotional stimuli generally trigger more EBA activity than neutral stimuli (Grèzes et al., 2007; Kret et al., submitted; Peelen, Atkinson, Andersson, & Vuilleumier, 2007). The role of the EBA in

emotional processing has not been fully understood yet and it is too early to claim that EBA is specifically sensitive for bodily features and less to the configural representation of a body such as the FG. Moreover, some interesting effects of gender were observed in this area. EBA is more active for bodily than facial expressions, especially when threatening, even more so when expressed by a male actor and above all when observed by a male participant.

Bodies may be processed without attention and visual awareness

There is an extensive literature on implicit (i.e., nonconscious) processing of facial expressions (de Gelder, De Haan, & Heywood, 2001; Eastwood & Smilek, 2005; Johnson, 2005; Vuilleumier, 2005). Nonconscious processing may occur either because attention is engaged elsewhere, so that the unattended stimulus also goes undetected (Mack, Pappas, Silverman, & Gay, 2002), or because conscious vision is directly prevented (Macknik & Livingstone, 1998; Weiskrantz, 2009). Research showing that faces and facial expressions are still processed under conditions of limited attention and awareness has contributed significantly to the view that faces have a special status. Recent evidence now shows a similar situation for bodily expressions. Two separate lines of evidence are provided by studies of neurological patients with attention disorders or with cortical blindness.

Patients with hemi-spatial neglect following lesions to the right parietal cortex fail to report leftward stimuli because of a deficit to orient their attention toward the left (contra-lesional) side of the space (Rafal, Egly, & Rhodes, 1994). It has been previously shown that emotional facial expressions presented to the left side tend to call for attention and are more often detected than neutral faces or objects, therefore partially overcoming the attentional bias (Tamietto, Geminiani, & de Gelder, 2005; Vuilleumier, 2002; Vuilleumier & Schwartz, 2001). We have replicated this effect by showing that also fearful bodily expressions may automatically summon spatial attention toward the neglected side, even under more stringent testing conditions than those previously used with faces (Tamietto, Geminiani, Genero, & de Gelder, 2007). Some differences, however, may be found in the neural correlates mediating conscious vs. nonconscious processing of faces and bodies in neglect patients. In fact, whereas in the case of faces the involvement of a fronto-limbic system seems relevant to recall attention and restore stimulus awareness (Vuilleumier, 2002), in the case of bodies

sensory-motor and interoceptive areas (e.g., insula) appear more critical (Tamietto & de Gelder, 2008).

Possibly the clearest example of processing bodily expression in the absence of stimulus awareness is provided by patients with damage to the visual cortex, because they literally cannot see the stimuli presented to the blind portion of their visual field. It is now established that under appropriate testing conditions such patients may reliably discriminate the visual properties of stimuli whose presence they are unaware (blindsight) (Weiskrantz, 2009). This ability to discriminate the attributes of unseen visual stimuli also extends to their affective valence (affective blindsight) (de Gelder & Tamietto, 2007). The initial reports used facial expressions and affective pictures, with positive results for the former stimuli and negative results for the latter, therefore suggesting a special status for faces in conveying nonconscious emotional information via subcortical pathways (Anders et al., 2004; de Gelder, Morris, & Dolan, 2005; de Gelder, Pourtois, van Raamsdonk, Vroomen, & Weiskrantz, 2001; de Gelder, Pourtois, & Weiskrantz, 2002; de Gelder, Vroomen, Pourtois, & Weiskrantz, 1999; Pegna, Khateb, Lazeyras, & Seghier, 2005). However, affective blindsight exists also for bodily expressions (de Gelder & Hadjikhani, 2006; Tamietto, Castelli et al., 2009; Tamietto & de Gelder, 2008). A comparison with bodies and faces is based on behavioral performance (i.e., response accuracy and latency) (de Gelder, Pourtois, & Hadjikhani, 2001; de Gelder et al., 1999; Tamietto & de Gelder, 2008), peripheral responses (e.g., expressive or autonomic reactions to the stimuli) (Tamietto, Castelli et al., 2009), or neural underpinnings of nonconscious processing (Anders et al., 2004; de Gelder & Hadjikhani, 2006; de Gelder et al., 2002; Morris, de Gelder, Weiskrantz, & Dolan, 2001; Pegna et al., 2005; Tamietto, Cauda et al., 2009). Analysis of all these response parameters in patients with affective blindsight indicates that faces and bodies induce highly similar responses. Therefore, the emerging picture shows that nonconscious processing of emotions is not specific for faces but it is clearly documented also for bodies. This indicates that the two types of stimuli share a common representation, possibly in subcortical structures such as the superior colliculus and amygdala responsible for coarse evaluation of the affective relevance of visual stimuli.

Finally, an interesting issue is whether emotional meaning may be extracted nonconsciously from specific parts of the face or body, or whether this requires prior analysis of the overall

configuration of the stimuli. There is initial evidence that specific regions of the face such as the eyes, may be sufficient in conveying the affective information outside visual awareness (Whalen et al., 2004). A recent study recording eye movements has found which specific body parts (hands, arms, legs and the position of the trunk) attracts the gaze that varies with the specific emotion displayed. For example, when observing angry bodily expressions, subjects fixate primarily the hands while for sad expressions they look at the face (Fridin, Barliya, Schechtman, de Gelder, & Flash, submitted).

Learning from lesions

An argument in favor of the higher overall category specificity faces than of bodies may be that of a specialized neurofunctional substrate for faces but not for bodies. We have already reviewed this argument as far as the neurologically intact brain is concerned. But a strong impetus for face and more generally category specificity of neural substrates comes from neuropsychological reports of patients with brain damage acquired in adulthood. There is a well known neuropsychological deficit related to impaired face recognition, labeled prosopagnosia or face blindness. These patients are impaired in recognizing faces, and very often have no recognition at all of an individual by the face only. Brain damage occurring in the normally developed brain that affects face perception is often localized in occipitotemporal cortex and temporal cortex (mid FG and IOG) unilaterally or bilaterally. The developmental counterpart of acquired prosopagnosia is now also increasingly reported. There is substantial similarity between acquired and developmental prosopagnosia at the behavioral level but there are many other differences (see de Gelder & Rouw, 2000 for a comparison).

It is important to specify the exact nature of the face disorder and this is still a matter of debate. The short definition of prosopagnosia characterizes it as a deficit in face recognition. But this is too broad and also too a-specific. We are in fact dealing with a deficit that affects recognition of personal identity from the sight of the face. Other dimensions of face information are processed mostly normally, like emotional expression, visual speech or gender. In fact, the reliable means of defining the typical face deficit of prosopagnosics is by establishing that there exists dissociation between the different dimensions of face perception,

some of which are impaired while others are intact. In addition, a dissociation must be established requiring evidence that the perception and recognition impairment is not present for non face stimuli. But there are to date only very few cases of pure prosopagnosia, where the perception and recognition deficit is restricted to the face and does not affect other object categories. To establish the presence of developmental prosopagnosia (DP) the same dimensions of dissociation need to be assessed. Thus when making the parallel argument for 'body blindness' the aspect of body perception that is impaired vs. the ones that are intact require evidence. These two dimensions of dissociation need to be taken into account. Therefore the equivalent of face deficit in the case of body perception is also likely to consist of one dimension of body perception and recognition.

On the other hand, using body stimuli offers a chance to advance the debate on category specificity of faces. For example, the behavioral and neuro-functional similarities between perceiving faces and bodies in neurologically intact observers raise the issue how bodies are processed in DP. We investigated perception of emotional and neutral faces and bodies in DPs and normal controls (Van den Stock, van de Riet, Righart, & de Gelder, 2008). The results showed a lower activation for neutral faces in FFA in the DP-group compared to the control group, but there was no difference between both groups for the emotional faces in FFA. We also compared activation of body expressions in face selective regions and of facial expressions in body selective regions. Our findings indicate that perceiving neutral faces results in a higher activation of EBA in the DP group, compared to the control group. Combined with the lower activation in FFA for neutral faces, this increased activation in EBA may indicate an anomalous processing route in the brains of DPs. It may be the case that (neutral) faces are processed in the areas more dominantly dedicated to body perception. On the other hand, we found a higher activation for perceiving bodies in IOG. These combined findings indicate that the neural correlates of perceiving faces and bodies in IOG and EBA show a lower degree of specificity in DP. These results are clearly tentative. Further research is needed to determine whether these differences in face and body processing between DP and controls reflect differences in processing routes or may result from compensation for the face processing deficits of DP.

Recently, a triple dissociation was demonstrated between faces, bodies and objects when different regions of the extrastriate occipital cortex were temporarily inactivated by transcranial magnetic stimulation (TMS) (Pitcher, Charles, Devlin, Walsh, & Duchaine, 2009). Stimulating EBA resulted in a selective body-processing deficit, whereas stimulating OFA and LOC (lateral occipital complex) (Malach et al., 1995) resulted in a selective impairment of face and object discrimination respectively. These findings strongly support a (partly) modular organization of the human brain. But here also an important issue is the whether this picture remains the same when functional aspects like for example the emotional expression are taken into account.

Models of body perception

In the last 5 years a few theoretical models of body perception have been advanced. The first one in de Gelder et al. (2004) was based on the only whole brain fMRI data available at that time and systematized the brain areas that differentially activated to neutral body actions and bodily expressions. It distinguishes between processes involved in low level body detection on the one hand and body perception and recognition and provides room for contribution of subcortical structures to detection and expression recognition. Integration of those initial results with new information provided by other techniques as well as by lesion studies led to the model in de Gelder (2006) in which body perception is envisaged along three interconnected networks: a predominantly subcortical one sustaining reflexive reactions, a cortical one sustaining visual analysis and reflective action and an interface system of bodily awareness.

Models addressing a more specific range of data provided by studies of neutral still bodies and focussing on the issue of part vs whole processes in body perception are provided in Taylor et al. (2007), Urgesi et al. (2007) and Hodzic et al. (2009). These models of neutral body perception are reminiscent of the earlier models of face processing in which typically a number of separate stages were distinguished. They tend to be hierarchical and serial and address functional and neural questions about face perception mainly from the vantage point of how recognition of personal identity is achieved. But as convergence grows between the researchers of face recognition in this narrow sense and those working on facial expression

recognition a rapprochement is seen between the two kinds of models. A major impetus for this rapprochement came from findings that facial expressions were perceived “earlier” than encoding of the face structure and that there was residual face processing ability in patients with cortical damage. These findings and others led to extended models of face processing encompassing both early and late processes, both expression and identity (Adolphs, 2002; de Gelder, Frissen, Barton, & Hadjikhani, 2003; de Gelder & Rouw, 2000; Haxby, Hoffman, & Gobbini, 2000) and involving conscious as well as unconscious, cortical but also subcortical structures and detection and recognition routes (de Gelder et al., 2003; de Gelder & Rouw, 2000).

Two sets of results available already about body perception allow us to anticipate that something similar is likely to be needed for bodies and for bodily expressions. One is that bodily expressions seem to trigger earlier activation in striate and extrastriate visual cortex, anterior temporal areas, but also more dorsal structures like parieto-occipital sulcus and intraparietal sulcus (Meeren et al., in preparation). Consistent with this and similar to findings about facial expressions, bodily expressions already show an emotional action trigger at the P1 component in the 100-120 window (Meeren et al., 2005; van Heijnsbergen et al., 2007), before the structural encoding of the stimulus has taken place. Importantly, those early activations testify to configural processing of the body stimulus as reflected in the fact that there is an early inversion effect (Meeren et al., 2008). The second important set of findings is the existence of residual processing of body stimuli in patients with cortical damage (de Gelder & Hadjikhani, 2006; Tamietto, Castelli et al., 2009; Tamietto & de Gelder, 2008) which is again similar to face processing, calling for body perception models that also acknowledge the contribution from subcortical structures which are important for rapid visual analysis and reactive behavior.

Of course, all models are tentative till more research is available. As our overview illustrates, a focus on isolated, single function areas is difficult to maintain when issues of control conditions and task demands are acknowledged. For example, when considering the function of EBA in itself, one issue is puzzling though. Urgesi et al. (2007) as well as Taylor et al. (2007) and Hodzich et al. (2009) attribute featural but not configural processing to EBA. Yet it emerges from studies of bodily expressions mentioned already (Grèzes et al., 2007; Pichon

et al., 2008) that EBA is sensitive to whether there is affective information in the body stimulus. This modulation by emotion may be compatible with EBA as a feature processor, in which case one would need to investigate which specific body part conveys the affective information. Alternatively, EBA does process the configuration of the stimulus after all. This alternative is consistent with our findings that EBA is differentially sensitive to affective information in the body (Grèzes et al., 2007; Pichon et al., 2008) when videos are used. This ambiguity as well as that concerning FG suggests that the role of an area is best investigated in connection with its role in a network.

Conclusion

For at least three decades faces occupied the most prominent position on the research agenda of psychologists, neuropsychologists, neurophysiologists and cognitive neuroscientists. The reasons for this situation are diverse. They range from recognition of the evolutionary and communicative importance of the face to conceptual and philosophical considerations about the need for a modular organization of the brain reflected in specialized perceptual abilities. Modular theorists have typically used the case of language and that of face perception as the clearest examples. In this paper we have tried to put current findings in perspective without however pushing them in the framework of the traditional question of modularity. We have argued that a number of different dimensions need to be considered when comparing the behavioral, neurofunctional and neuropsychological basis of faces and bodies. The debate is open, but it would be a pity to limit it to the traditional questions and debates about specialized modules and categorical representation in the brain. When viewed in a broader perspective, faces and bodies are comparable because they both convey information that is essential for social interaction. Yet they each fulfill this role in a different way. It seems likely that faces are used for fine-grained analysis of communication intention and possibly also convey a broader range of subtle emotions in a setting of close-by person to person. In contrast, while body language is still important for close-by interaction, bodily expression allow for appraisal of action intention and emotions over larger distances. These differences are lost when one focuses only representations issues.

References

- Adolphs, R. (2002). Neural systems for recognizing emotion. *Curr Opin Neurobiol*, 12(2), 169-177.
- Anders, S., Birbaumer, N., Sadowski, B., Erb, M., Mader, I., Grodd, W., et al. (2004). Parietal somatosensory association cortex mediates affective blindsight. *Nat Neurosci*, 7(4), 339-340.
- Astafiev, S. V., Stanley, C. M., Shulman, G. L., & Corbetta, M. (2004). Extrastriate body area in human occipital cortex responds to the performance of motor actions. *Nat Neurosci*, 7(5), 542-548.
- Atkinson, A. P., Dittrich, W. H., Gemmell, A. J., & Young, A. W. (2004). Emotion perception from dynamic and static body expressions in point-light and full-light displays. *Perception*, 33(6), 717-746.
- Barracough, N. E., Xiao, D., Oram, M. W., & Perrett, D. I. (2006). The sensitivity of primate STS neurons to walking sequences and to the degree of articulation in static images. *Prog Brain Res*, 154, 135-148.
- Bentin, S., Allison, T., Puce, A., Perez, E., & et al. (1996). Electrophysiological studies of face perception in humans. *Journal of Cognitive Neuroscience*, 8(6), 551-565.
- Breiter, H. C., Etcoff, N. L., Whalen, P. J., Kennedy, W. A., Rauch, S. L., Buckner, R. L., et al. (1996). Response and Habituation Of the Human Amygdala During Visual Processing Of Facial Expression. *Neuron*, 17(5), 875-887.
- Chan, A. W., Peelen, M. V., & Downing, P. E. (2004). The effect of viewpoint on body representation in the extrastriate body area. *Neuroreport*, 15(15), 2407-2410.
- Cox, D., Meyers, E., & Sinha, P. (2004). Contextually evoked object-specific responses in human visual cortex. *Science*, 304(5667), 115-117.
- de Gelder, B. (2006). Towards the neurobiology of emotional body language. *Nature Reviews Neuroscience*, 7(3), 242-249.
- de Gelder, B. (2009). Why bodies? Twelve reasons for including bodily expressions in affective neuroscience. *Proc R Soc Lond B Biol Sci*, (in press).

- de Gelder, B., De Haan, E. H. F., & Heywood, C. A. (Eds.). (2001). *Out of mind: Varieties of unconscious processes*. London: Oxford University Press.
- de Gelder, B., Frissen, I., Barton, J., & Hadjikhani, N. (2003). A modulatory role for facial expressions in prosopagnosia. *Proc Natl Acad Sci U S A*, 100(22), 13105-13110.
- de Gelder, B., & Hadjikhani, N. (2006). Non-conscious recognition of emotional body language. *Neuroreport*, 17(6), 583-586.
- de Gelder, B., Morris, J. S., & Dolan, R. J. (2005). Unconscious fear influences emotional awareness of faces and voices. *Proc Natl Acad Sci U S A*, 102(51), 18682-18687.
- de Gelder, B., & Partan, S. (2009). The neural basis of perceiving emotional bodily expressions in monkeys. *Neuroreport*, 20(7), 642-646.
- de Gelder, B., Pourtois, G., & Hadjikhani, N. (2001). The time course of face recognition: learning from face evidence for a dual route model. *International Journal of Psychophysiology*, 41(3), 215-215.
- de Gelder, B., Pourtois, G., van Raamsdonk, M., Vroomen, J., & Weiskrantz, L. (2001). Unseen stimuli modulate conscious visual experience: evidence from inter-hemispheric summation. *Neuroreport*, 12(2), 385-391.
- de Gelder, B., Pourtois, G., & Weiskrantz, L. (2002). Fear recognition in the voice is modulated by unconsciously recognized facial expressions but not by unconsciously recognized affective pictures. *Proc Natl Acad Sci U S A*, 99(6), 4121-4126.
- de Gelder, B., & Rouw, R. (2000). Configural face processes in acquired and developmental prosopagnosia: evidence for two separate face systems? *Neuroreport*, 11(14), 3145-3150.
- de Gelder, B., Snyder, J., Greve, D., Gerard, G., & Hadjikhani, N. (2004). Fear fosters flight: A mechanism for fear contagion when perceiving emotion expressed by a whole body. *Proc Natl Acad Sci U S A*, 101(47), 16701-16706.
- de Gelder, B., & Tamietto, M. (2007). Affective blindsight. *Scholarpedia*, 2(10), 3555.
- de Gelder, B., Vroomen, J., Pourtois, G., & Weiskrantz, L. (1999). Non-conscious recognition of affect in the absence of striate cortex. *Neuroreport*, 10(18), 3759-3763.

- de Meijer, M. (1989). The contribution of general features of body movement to the attribution of emotions. *J Nonverbal Behavior*, 13, 247-268.
- Desimone, R., Albright, T. D., Gross, C. G., & Bruce, C. (1984). Stimulus-selective properties of inferior temporal neurons in the macaque. *J Neurosci*, 4(8), 2051-2062.
- Downing, P. E., Chan, A. W., Peelen, M. V., Dodds, C. M., & Kanwisher, N. (2006). Domain specificity in visual cortex. *Cereb Cortex*, 16(10), 1453-1461.
- Downing, P. E., Jiang, Y., Shuman, M., & Kanwisher, N. (2001). A cortical area selective for visual processing of the human body. *Science*, 293(5539), 2470-2473.
- Downing, P. E., Peelen, M. V., Wiggett, A. J., & Tew, B. D. (2006). The role of the extrastriate body area in action perception. *Soc Neurosci*, 1(1), 52-62.
- Downing, P. E., Wiggett, A. J., & Peelen, M. V. (2007). Functional magnetic resonance imaging investigation of overlapping lateral occipitotemporal activations using multi-voxel pattern analysis. *J Neurosci*, 27(1), 226-233.
- Eastwood, J. D., & Smilek, D. (2005). Functional consequences of perceiving facial expressions of emotion without awareness. *Conscious Cogn*, 14(3), 565-584.
- Eimer, M. (2000). The face-specific N170 component reflects late stages in the structural encoding of faces. *Neuroreport*, 11(10), 2319-2324.
- Farah, M. (1990). *Visual agnosia: Disorders of visual recognition and what they tell us about normal vision*. Cambridge: MIT Press.
- Fridin, M., Barliya, A., Schechtman, E., de Gelder, B., & Flash, T. (submitted). Gaze and the perception of emotional body language (EBL). An eye tracking study.
- Gliga, T., & Dehaene-Lambertz, G. (2005). Structural encoding of body and face in human infants and adults. *Journal of Cognitive Neuroscience*, 17, 1328-1340.
- Grèzes, J., & Decety, J. (2001). Functional anatomy of execution, mental simulation, observation, and verb generation of actions: a meta-analysis. *Hum Brain Mapp*, 12(1), 1-19.
- Grèzes, J., Pichon, S., & de Gelder, B. (2007). Perceiving fear in dynamic body expressions. *Neuroimage*, 35(2), 959-967.

- Gross, C. G., Bender, D. B., & Rocha-Miranda, C. E. (1969). Visual receptive fields of neurons in inferotemporal cortex of the monkey. *Science*, 166(910), 1303-1306.
- Gross, C. G., Rocha-Miranda, C. E., & Bender, D. B. (1972). Visual properties of neurons in inferotemporal cortex of the Macaque. *J Neurophysiol*, 35(1), 96-111.
- Hadjikhani, N., & de Gelder, B. (2003). Seeing fearful body expressions activates the fusiform cortex and amygdala. *Curr Biol*, 13(24), 2201-2205.
- Haxby, J. V., Hoffman, E. A., & Gobbini, M. I. (2000). The distributed human neural system for face perception. *Trends Cogn Sci*, 4(6), 223-233.
- Hodzic, A., Kaas, A., Muckli, L., Stirn, A., & Singer, W. (2009). Distinct cortical networks for the detection and identification of human body. *Neuroimage*, 45(4), 1264-1271.
- Jellema, T., & Perrett, D. I. (2003a). Cells in monkey STS responsive to articulated body motions and consequent static posture: a case of implied motion? *Neuropsychologia*, 41(13), 1728-1737.
- Jellema, T., & Perrett, D. I. (2003b). Perceptual history influences neural responses to face and body postures. *J Cogn Neurosci*, 15(7), 961-971.
- Johnson, M. H. (2005). Subcortical face processing. *Nat Rev Neurosci*, 6(10), 766-774.
- Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: a module in human extrastriate cortex specialized for face perception. *J Neurosci*, 17(11), 4302-4311.
- Kiani, R., Esteky, H., Mirpour, K., & Tanaka, K. (2007). Object category structure in response patterns of neuronal population in monkey inferior temporal cortex. *J Neurophysiol*, 97(6), 4296-4309.
- Kret, M. E., Pichon, S., Grèzes, J., & de Gelder, B. (submitted). Gender specific brain activations in perceiving threat from dynamic faces and bodies.
- Kriegeskorte, N., Mur, M., Ruff, D. A., Kiani, R., Bodurka, J., Esteky, H., et al. (2008). Matching categorical object representations in inferior temporal cortex of man and monkey. *Neuron*, 60(6), 1126-1141.

- Mack, A., Pappas, Z., Silverman, M., & Gay, R. (2002). What we see: inattention and the capture of attention by meaning. *Conscious Cogn*, 11(4), 488-506.
- Macknik, S. L., & Livingstone, M. S. (1998). Neuronal correlates of visibility and invisibility in the primate visual system. *Nat Neurosci*, 1(2), 144-149.
- Malach, R., Reppas, J. B., Benson, R. R., Kwong, K. K., Jiang, H., Kennedy, W. A., et al. (1995). Object-related activity revealed by functional magnetic resonance imaging in human occipital cortex. *Proc Natl Acad Sci U S A*, 92(18), 8135-8139.
- McCarthy, G., Puce, A., Belger, A., & Allison, T. (1999). Electrophysiological studies of human face perception. II: Response properties of face-specific potentials generated in occipitotemporal cortex. *Cereb Cortex*, 9(5), 431-444.
- Meeren, H. K., Hadjikhani, N., Ahlfors, S. P., Hamalainen, M. S., & de Gelder, B. (2008). Early category-specific cortical activation revealed by visual stimulus inversion. *PLoS ONE*, 3(10), e3503.
- Meeren, H. K., van Heijnsbergen, C. C., & de Gelder, B. (2005). Rapid perceptual integration of facial expression and emotional body language. *Proc Natl Acad Sci U S A*, 102(45), 16518-16523.
- Minnebusch, D. A., & Daum, I. (2009). Neuropsychological mechanisms of visual face and body perception. *Neurosci Biobehav Rev*.
- Morris, J. S., de Gelder, B., Weiskrantz, L., & Dolan, R. J. (2001). Differential extrageniculostriate and amygdala responses to presentation of emotional faces in a cortically blind field. *Brain*, 124(Pt 6), 1241-1252.
- Oram, M. W., & Perrett, D. I. (1996). Integration of form and motion in the anterior superior temporal polysensory area (STPa) of the macaque monkey. *J Neurophysiol*, 76(1), 109-129.
- Peelen, M. V., Atkinson, A. P., Andersson, F., & Vuilleumier, P. (2007). Emotional modulation of body-selective visual areas. *Social, Cognitive and Affective Neurosciences*, 2, 274-283.
- Peelen, M. V., & Downing, P. E. (2005a). Is the extrastriate body area involved in motor actions? *Nat Neurosci*, 8(2), 125; author reply 125-126.

- Peelen, M. V., & Downing, P. E. (2005b). Selectivity for the human body in the fusiform gyrus. *J Neurophysiol*, 93(1), 603-608.
- Peelen, M. V., Wiggett, A. J., & Downing, P. E. (2006). Patterns of fMRI activity dissociate overlapping functional brain areas that respond to biological motion. *Neuron*, 49(6), 815-822.
- Pegna, A. J., Khateb, A., Lazeyras, F., & Seghier, M. L. (2005). Discriminating emotional faces without primary visual cortices involves the right amygdala. *Nat Neurosci*, 8(1), 24-25.
- Perrett, D. I., Smith, P. A., Mistlin, A. J., Chitty, A. J., Head, A. S., Potter, D. D., et al. (1985). Visual analysis of body movements by neurones in the temporal cortex of the macaque monkey: a preliminary report. *Behav Brain Res*, 16(2-3), 153-170.
- Pichon, S., de Gelder, B., & Grezes, J. (2008). Emotional modulation of visual and motor areas by dynamic body expressions of anger. *Social Neuroscience*, 3(3-4), 199-212.
- Pinsk, M. A., DeSimone, K., Moore, T., Gross, C. G., & Kastner, S. (2005). Representations of faces and body parts in macaque temporal cortex: a functional MRI study. *Proc Natl Acad Sci U S A*, 102(19), 6996-7001.
- Pitcher, D., Charles, L., Devlin, J. T., Walsh, V., & Duchaine, B. (2009). Triple dissociation of faces, bodies, and objects in extrastriate cortex. *Curr Biol*, 19(4), 319-324.
- Pourtois, G., Peelen, M. V., Spinelli, L., Seeck, M., & Vuilleumier, P. (2007). Direct intracranial recording of body-selective responses in human extrastriate visual cortex. *Neuropsychologia*, 45, 2621-2625.
- Rafal, R., Egly, R., & Rhodes, D. (1994). Effects of inhibition of return on voluntary and visually guided saccades. *Can J Exp Psychol*, 48(2), 284-300.
- Reed, C. L., Stone, V. E., Bozova, S., & Tanaka, J. (2003). The body-inversion effect. *Psychol Sci*, 14(4), 302-308.
- Righart, R., & de Gelder, B. (2007). Impaired face and body perception in developmental prosopagnosia. *Proceedings of the National Academy of Sciences*, 104(43), 17234-17238.

- Righart, R., & de Gelder, B. (2008). Facial expressions and emotional scene gist. *cognitive, Affective, & Behavioural Neuroscience*, 8(3), 264-272.
- Schwarzlose, R. F., Baker, C. I., & Kanwisher, N. (2005). Separate face and body selectivity on the fusiform gyrus. *J Neurosci*, 25(47), 11055-11059.
- Sinke, C. B. A., Sorger, B., Goebel, R., & de Gelder, B. (2009). Is it just a tease or a real threat? Judging social interactions from bodily expressions. *NeuroImage*, (in press).
- Spiridon, M., Fischl, B., & Kanwisher, N. (2006). Location and spatial profile of category-specific regions in human extrastriate cortex. *Hum Brain Mapp*, 27(1), 77-89.
- Stekelenburg, J. J., & de Gelder, B. (2004). The neural correlates of perceiving human bodies: an ERP study on the body-inversion effect. *Neuroreport*, 15(5), 777-780.
- Tamietto, M., Castelli, L., Vighetti, S., Perozzo, P., Geminiani, G., Weiskrantz, L., et al. (2009). Unseen facial and bodily expressions trigger fast emotional reactions. *Proc Natl Acad Sci U S A*.
- Tamietto, M., Cauda, F., Corazzini, L. L., Savazzi, S., Marzi, C. A., Goebel, R., et al. (2009). Collicular Vision Guides Nonconscious Behavior. *Journal of Cognitive Neuroscience*.
- Tamietto, M., & de Gelder, B. (2008). Affective blindsight in the intact brain: neural interhemispheric summation for unseen fearful expressions. *Neuropsychologia*, 46(3), 820-828.
- Tamietto, M., Geminiani, G., & de Gelder, B. (2005). *Inter-hemispheric interaction for bodily emotional expressions: Is the right hemisphere superiority related to facial rather than emotional processing?* Paper presented at the Proceedings of the XVIII European Conference on Visual Perception Annual Meeting, A Coruña, Spain, , August 22-26.
- Tamietto, M., Geminiani, G., Genero, R., & de Gelder, B. (2007). Seeing fearful body language overcomes attentional deficits in patients with neglect. *Journal of Cognitive Neuroscience*, 19(3), 445-454.
- Taylor, J. C., Wiggett, A. J., & Downing, P. E. (2007). Functional MRI analysis of body and body part representations in the extrastriate and fusiform body areas. *Journal of Neurophysiology*, 98(3), 1626-1633.

- Thierry, G., Pegna, A. J., Dodds, C., Roberts, M., Basan, S., & Downing, P. (2006). An event-related potential component sensitive to images of the human body. *Neuroimage*, 32(2), 871-879.
- Tsao, D. Y., Freiwald, W. A., Knutsen, T. A., Mandeville, J. B., & Tootell, R. B. (2003). Faces and objects in macaque cerebral cortex. *Nat Neurosci*, 6(9), 989-995.
- Tsao, D. Y., Freiwald, W. A., Tootell, R. B., & Livingstone, M. S. (2006). A cortical region consisting entirely of face-selective cells. *Science*, 311(5761), 670-674.
- Urgesi, C., Calvo-Merino, B., Haggard, P., & Aglioti, S. M. (2007). Transcranial magnetic stimulation reveals two cortical pathways for visual body processing. *J Neurosci*, 27(30), 8023-8030.
- van de Riet, W. A., Grezes, J., & de Gelder, B. (2009). Specific and common brain regions involved in the perception of faces and bodies and the representation of their emotional expressions. *Social Neuroscience*, 4(2), 101-120.
- Van den Stock, J., van de Riet, W. A., Righart, R., & de Gelder, B. (2008). Neural correlates of perceiving emotional faces and bodies in developmental prosopagnosia: an event-related fMRI-study. *PLoS ONE*, 3(9), e3195.
- van Heijnsbergen, C. C., Meeren, H. K., Grezes, J., & de Gelder, B. (2007). Rapid detection of fear in body expressions, an ERP study. *Brain Res*, 1186, 233-241.
- Vuilleumier, P. (2002). Perceived gaze direction in faces and spatial attention: a study in patients with parietal damage and unilateral neglect. *Neuropsychologia*, 40(7), 1013-1026.
- Vuilleumier, P. (2005). How brains beware: neural mechanisms of emotional attention. *Trends Cogn Sci*, 9(12), 585-594.
- Vuilleumier, P., Armony, J. L., Driver, J., & Dolan, R. J. (2001). Effects of attention and emotion on face processing in the human brain: an event-related fMRI study. *Neuron*, 30(3), 829-841.
- Vuilleumier, P., & Schwartz, S. (2001). Beware and be aware: capture of spatial attention by fear-related stimuli in neglect. *Neuroreport*, 12(6), 1119-1122.

- Wachsmuth, E., Oram, M. W., & Perrett, D. I. (1994). Recognition of objects and their component parts: responses of single units in the temporal cortex of the macaque. *Cereb Cortex*, 4(5), 509-522.
- Weiskrantz, L. (2009). *Blindsight: a case study spanning 35 years and new developments*. Oxford: Oxford University Press.
- Whalen, P. J., Kagan, J., Cook, R. G., Davis, F. C., Kim, H., Polis, S., et al. (2004). Human amygdala responsivity to masked fearful eye whites. *Science*, 306(5704), 2061.
- Yin, R. K. (1969). Looking at upside-down faces. *Journal of Experimental Psychology*, 81, 141-145.

Chapter 3: Seeing bodily expressions in context: contagion between bodily expression and scene³

Introduction

In everyday life we interact with people as we encounter them in their natural environment and react to their facial and body expression. Observers are capable of reacting rapidly to the gist of a scene (Thorpe, Fize, & Marlot, 1996). Similarly, affective information from body postures is processed rapidly (B. de Gelder et al., 2009). There is now also evidence from behavioral studies and from EEG recordings that observers attending to the central face stimulus or the central body stimulus still process the affective gist of the surrounding scene. For example, the face sensitive N170 is modified by the presence of a scene that is affectively congruent with the facial expression (Righart & de Gelder, 2006). Similarly, observers are more accurate in rapidly categorizing a bodily expression when it is affectively congruent with the affective gist of the surrounding scene (Kret & de Gelder, submitted).

In this study we used magnetic resonance imaging (fMRI), to investigate the neural basis of processing bodily expressions presented against the background of a natural scene. We combined fearful and neutral bodies with threatening and neutral scenes to create realistically looking compound stimuli and focused on the known category sensitive brain areas for bodies (the so-called EBA: extrastriate body area (Downing, Jiang, Shuman, & Kanwisher, 2001) and FBA: fusiform body area (Peelen & Downing, 2005)) and scenes (the PPA: parahippocampal place area (R. Epstein & Kanwisher, 1998) and RSC: retrosplenial complex (Bar & Aminoff, 2003)).

Method

Subjects. Twenty-one subjects (7 males; mean age (SD): 25.14 (9.5)), took part in the fMRI-experiment. All subjects had normal or corrected-to-normal vision and no neurological or

³ This chapter is submitted.

psychiatric history. Informed written consent was obtained in accordance with the declaration of Helsinki.

Task and procedures.

a) Body-scene experiment

Pictures of whole body expressions with the faces blurred were taken from our own validated database (Van den Stock, Righart, & de Gelder, 2007). Twenty-four images (12 identities, each with a neutral and a fearful expression) were selected for use in the present study (all recognized accurately above 80% in the pilot study). Scenes of happy, threatening, neutral, sad or disgusting everyday situation were downloaded from the internet. We selected three familiar scene categories (houses, cars and landscapes) that involve the same objects but with the different affective significances. For example, a house on fire or a holiday cottage, a damaged car in an accident or a shiny new convertible. None of the scenes displayed humans or animals. In a pilot study the scenes were presented one by one for 4000 ms with a 4000 ms interstimulus interval. Participants were instructed to categorize as accurately and as quickly as possible the pictures according to the emotion they induced in the observer. Based on the results we selected 24 scenes (12 threatening and 12 neutral) for the present experiment (all recognized correctly above 70%). Each category contained 4 exemplars with a car, 4 with a building and 4 with a landscape. We created scrambled versions of every scene, by dividing it in 10 000 (100 x 100) squares and randomly rearranging the squares. Each of the 24 scenes was combined into a compound stimulus in four different combinations (with a fearful male body, a fearful female body, a neutral male body and a neutral female body) resulting in 96 realistically looking compound stimuli. The control stimuli consisted of scrambled scenes with natural bodies and of natural scenes with triangles instead of bodies. The 24 scrambled scenes were combined once with a fearful body and once with a neutral body, leading to an additional 48 compound stimuli. We also paired every scene with a white triangle (24 intact scenes + 24 scrambled scenes). These scene-triangle combinations were used as a control condition instead of only the scenes in order to maintain the same task in all conditions (see below) and to ensure all stimuli had a clear foreground/background structure. This procedure results in 192 unique compound stimuli. (See Fig. 1 for examples.)

Stimuli were presented in blocks of 9000 ms, separated by fixation blocks of 15750 ms.

Within a block, eight stimuli were presented for 800 ms with an ISI of 370 ms, during which a

grey screen was shown. In fixation blocks, a gray screen with a black fixation cross was presented. We used a 3 (body: fearful, neutral and triangle) x 3 (scene: threatening, neutral and scrambled) factorial design. Participants were given an oddball detection task and instructed to press the response button when the foreground figure (body or triangle) was shown upside-down. A run lasted 711s and consisted of 31 experimental blocks and 32 fixation blocks. The order of the blocks was randomized. In 4 of the 31 blocks (13%) an oddball stimulus occurred, while the other 27 blocks consisted of 3 blocks of every condition. The experiment consisted of 4 runs.

b) object category localizer

Stimuli for the object category localizer consisted of images of neutral faces taken from the Karolinska Directed Emotional Faces set (Lundqvist, Flykt, & Öhman, 1998), neutral human body postures with the faces blurred that came from the Bodily Expressive Action Stimulus Test (BEAST) (B de Gelder & Van den Stock, 2008), buildings (courtesy of R. Goebel) and handheld tools (courtesy of M. Peelen). There were 20 exemplars of each category. Half of the face images and half of the body images were male.

Stimuli were presented in a blocked design. Stimulation blocks (12000 ms) were alternated with fixation blocks (14000 ms) on which a dark screen with white fixation cross was shown. Within a stimulation block, 12 stimuli were presented one by one with a duration of 450 ms and an inter stimulus interval of 600 ms, during which a gray screen was presented. The localizer run was pseudo-randomized with five consecutive series. A series contained one block of every stimulus category presented in a randomized order, resulting in five blocks per stimulus category. Half of the subjects were scanned with a parallel version in which the order of the stimulation blocks was reversed. The total duration of the run was 534 s (8.9 min). Participants performed a one-back task. The mean number of one back targets per block was 1 (range 0-2) with an equal distribution among conditions.

c) behavioral experiment

After the scanning session, 15 of the subjects (6 male; mean age (SD): 23 (4.0)) participated in a behavioral experiment. All body-scene stimuli were randomly presented twice one by one for 800 ms while subjects were instructed to categorize the emotion expressed by the

body in a two alternative forced choice task (fearful or neutral). No response time limited was given, but the instructions stated to respond as accurately and quickly as possible.

fMRI scan acquisition and data analysis. Brain imaging was performed on a Siemens MAGNETOM Allegra 3T MR head scanner at the Maastricht Brain Imaging Center, Maastricht University.

a) scan protocol and analysis

All participants underwent four experimental runs, in which 348 T2*-weighted BOLD contrast volumes were acquired. On the basis of structural information from a 9-slice localizer scan, 42 axial slices (slice thickness=2.5 mm; no gap; inplane resolution= 3.5x3.5mm; matrix size=64x64; FOV=224 mm) were positioned to cover the whole brain (TE=25; TR=2250 ms; flip angle=90°). Slices were scanned in an interleaved ascending order. A high-resolution T1-weighted anatomical image (voxel size=1x1x1 mm) was acquired for each subject using a three-dimensional magnetization-prepared rapid acquisition gradient echo (MP-RAGE) sequence (TR = 2250ms; TE = 2.6 ms; matrix size=256x256; 192 slices). Two functional runs were followed by the structural scan, after which the two remaining functional runs were completed.

The object category localizer run came at the end of the session. 267 T2*-weighted BOLD contrast volumes were acquired, consisting each of 28 axial slices (slice thickness=2 mm; no gap; inplane resolution= 2x2 mm; matrix size=128x128; FOV=256 mm) were positioned to cover the (ventral) occipito-temporal cortex (TE=30; TR=2000 ms; flip angle=90°). Slices were scanned in an interleaved ascending order. 267 functional volumes were acquired.

Imaging data were analysed using BrainVoyager QX software (Brain Innovation, Maastricht, Netherlands). Structural scans were segmented to delineate white matter from grey matter and we performed cortex-based inter-subject alignment based on the gyral/sulcal pattern of individual brains (Goebel, Esposito, & Formisano, 2006) to maximize anatomical between-subjects alignment. On the bases of this, an average cortical reconstruction was made of all individual brains. The first two volumes of every functional run were discarded to allow for

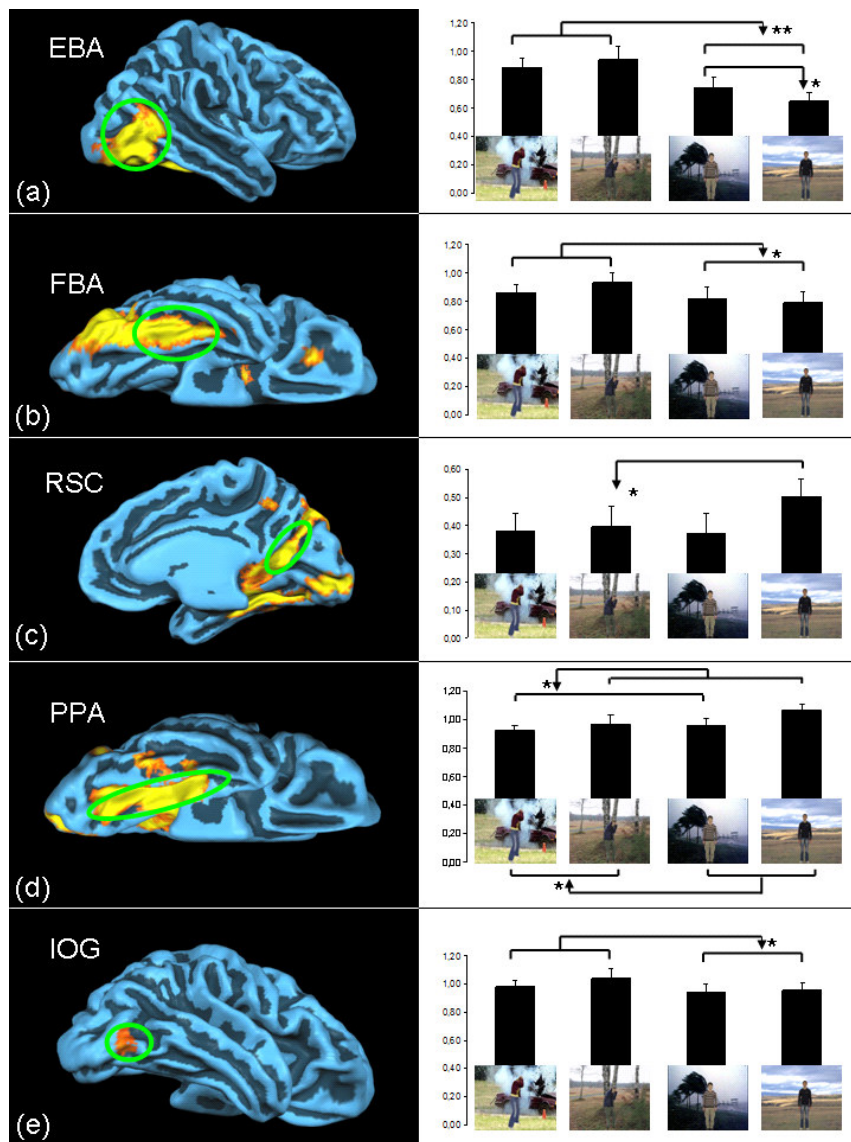
T1 equilibration. Preprocessing of the functional data included slice scan time correction (cubic spline interpolation), 3D motion correction (trilinear/sinc interpolation) and temporal filtering (high pass GLM-Fourier of 2 sines/cosines). Functional data were then co-registered with the anatomical volume and transferred into Talairach space (Talairach & Tournoux, 1988). Analysis of cortical activation included cortex based inter-subject alignment (Goebel et al., 2006). The statistical analysis was based on the General Linear Model, with each condition defined as a predictor plus one for the oddball.

Results and Discussion

We focused on cortical activation in the right hemisphere, since previous studies have shown that categorical selectivity is observed most clearly in the right hemisphere (Kanwisher, McDermott, & Chun, 1997; Peelen & Downing, 2005). Subcortical activation in both hemispheres was also inspected.

We explored the cortical areas that are sensitive to perception of both factors in a 3 (body: fearful, neutral or triangle) x 3 (scene: threatening, neutral or scramble) random effects ANOVA on the blood oxygenation level-dependent (BOLD) measurements ($P < 0.001$, corrected for multiple comparisons (Forman et al., 1995)). The resulting beta-values of the different clusters were normalized and submitted to a 3 (body) x 3 (scene) repeated measures ANOVA. Planned comparisons post hoc tests were administered to follow up on the main and interaction effects. Brain areas sensitive to body perception included a large L-shaped area in the occipito-temporal cortex (most of the fusiform gyrus), extending dorsally at the junction of the occipital and temporal lobe (see Fig1). This area comprises the well known body selective surfaces in the middle temporal and occipital gyrus (EBA (Downing et al., 2001)) and in the fusiform gyrus (FBA (Peelen & Downing, 2005)). To extract these areas from the large cluster on a subject-level, we defined in every subject EBA and FBA on the basis of an object category localizer (bodies vs tools and houses; $P < 0.05$, uncorrected). For EBA we selected the set of contiguous voxels at the junction of the middle temporal and middle occipital gyrus (TAL: $41 < x < 48$; $-75 < y < -54$; $-6 < z < 12$) and for FBA in the middle part of the fusiform gyrus (TAL: $26 < x < 46$; $-51 < y < -29$; $-17 < z < 12$). A full list of activation foci and post hoc analysis results is presented in Table 1.

The neural network sensitive to scene perception contained the well known parahippocampal place area (PPA) (R. Epstein & Kanwisher, 1998), and the retrosplenial cortex (RSC) (Bar & Aminoff, 2003). We localized these regions in every subject using the object category localizer by comparing buildings with tools, bodies and faces ($P < 0.05$). For PPA we selected the significant voxels in the parahippocampal gyrus (TAL: $19 < x < 28$; $-58 < y < -34$; $-12 < z < 6$) and for the RSC in the region of the parieto-occipital sulcus (TAL: $7 < x < 19$; $-62 < y < -36$; $-4 < z < 19$).



Left: Cortical activations of main and interaction effects following a 3 (body) X 3 (scene) ANOVA. (a) and (b): main effect of body, (c) and (d): main effect of scene and (e): body x scene interaction. Right: beta-values of the conditions with bodies and scenes. The values of (a) to (d) are extracted from subject based contrasts from localizer data (bodies vs tools and buildings for EBA (a) and FBA (b) and buildings vs tools, bodies and faces for RSC (c) and PPA (d). The beta-values for the interaction ROI (e) are extracted from the 3x3 group analysis. Conditions are represented on the X axis (from left to right: fearful body in threatening scene; fearful body in neutral scene; neutral body in threatening scene; neutral body in neutral scene). Error bars represent one standard error of the mean (SEM). * $P < 0.05$; ** $P < 0.001$.

Table 1. Areas sensitive to perception of bodily expressions, scenes and body x scene interaction, with results of post-hoc testing on normalized beta-values. Abbreviations EBA= extrastriate body area; FBA= fusiform body area; PHC= parahippocampal cortex; PPA= parahippocampal place area; RSC= retrosplenial complex; FG= fusiform gyrus; IOG= inferior occipital gyrus; Fb= Fearful body; Nb= Neutral body; t= triangle; Ts= threatening scene; Ns= neutral scene; Ss= scrambled scene. ¹: cortical subject-based defined area; ²: cortical group-based defined area; ³: whole brain group-based defined area. * $P < 0.001$; ** $P < 0.05$

	TAL-coordinates			Main effect of body	Main effect of scene	Interaction
	x	y	z			
Main effect of body						
EBA ¹	44	-65	4	Fb>Nb ^{**} Fb>t ^{**} Nb>t ^{**}	Ts>Ss ^{**} Ns>Ss ^{**}	NbTs>NbNs [*] NbTs>NbSs ^{**} NbNs>NbSs [*]
FBA ¹	40	-42	-14	Fb>Nb [*] Fb>t ^{**} Nb>t ^{**}	Ts>Ss ^{**} Ns>Ss ^{**}	NbTs>NbSs [*]
Orbitofrontal ²	27	31	-4	Fb>t ^{**} Nb>t ^{**}		
Cuneus ²	10	-89	12	Nb>Fb [*] Nb>t ^{**} Fb>t ^{**}		
Precuneus ²	7	-50	44	Nb>Fb [*] Nb>t ^{**} Fb>t [*]	Ts>Ss ^{**} Ns>Ss ^{**}	NbTs>NbSs [*] NbNs>NbSs [*]

Anterior PHC ²	28	-2	-14	Fb>t [*] Nb>t ^{**}		
Amygdala (right) ³	18	-1	-11	Nb>Fb [*] Nb>t ^{**} Fb>t ^{**}		
Main effect of scene						
PPA ¹	24	-43	-8	Nb>Fb [*]	Ns>Ts [*] Ts>Ss ^{**} Ns>Ss ^{**}	FbNs>INs [*]
RSC ¹	14	-51	10		Ts>Ss ^{**} Ns>Ss ^{**}	NbNs>FbNs [*]
Calcarine sulcus ²	9	-83	-1		Sc>Ts ^{**} Sc>Ns ^{**}	
Precuneus ²	7	-50	44	Nb>Fb [*] Fb>t [*] Nb>t ^{**}	Ts>Ss ^{**} Ns>Ss ^{**}	FbNs>INs [*] NbNs>INs ^{**}
Insula ²	37	-4	1		Sc>Ns ^{**} Sc>Ts [*]	NbNs>FbNs [*]
Anterior FG ²	36	-27	-17	Fb>t [*]	Ts>Ss ^{**}	FbTs>tTs [*]

			Nb>t ^{ns}	Ns>Ss ^{ns}	NbTs>tTs [*] FbNs>tNs [*] NbNs>tNs [*]
Posterior FG ²	33	-42	-13	Ts>Ss ^{ns} Ns>Ss ^{ns}	FbTs>tTs ^{ns} NbTs>tTs [*] FbNs>tNs [*] NbNs>tNs [*]
Caudate nucleus (right) ³	16	6	17	Sc>Ns ^{ns} Sc>Ts ^{ns}	
Body x Scene Interaction					
	IOG ²	42	-68	-5	NbTs>NbSs [*] FbTs>tTs ^{ns} NbTs>tTs ^{ns} FbNs>tNs ^{ns} NbNs>tNs ^{ns}

Fearful bodily expressions triggered higher activation levels than neutral bodies in EBA ($P < 0.0001$) and FBA ($P < 0.0134$) (see Fig1). Furthermore, EBA activation was higher for neutral bodies in threatening scenes than neutral bodies in neutral scenes ($P < 0.0335$). This affective scene effect in EBA was not found for fearful bodies or control stimuli (triangles) ($P < 0.2640$). The activity bias of neutral bodies in threatening scenes towards the activity level of fearful bodies is consistent with an emotional spillover effect from threatening scenes on neutral bodies. Behavioral data showed a compatible pattern: median reaction times (RT) showed a significantly higher RT for neutral bodies in a fearful context than for neutral bodies in a neutral context ($P < 0.0292$). A perceptual explanation may be that neutral bodies are emotionally ambiguous and that the emotional gist of the background scene colors the observers perception of the body stimulus in the foreground. This crossmodal bias effect reducing stimulus ambiguity is known from multisensory audiovisual perception studies (B. de Gelder & Bertelson, 2003) and obtains equally in the visual-visual domain.

Secondly, in PPA neutral scenes surprisingly elicited higher activity than threatening scenes ($P < 0.0167$). Research on PPA has centered on its role in spatial, navigational and contextual processing (Bar, 2004; R. A. Epstein, 2008) and little is known about its response to emotional scenes. Our data show that neutral scenes trigger more activity than threatening ones. This may indicate that threatening scenes trigger less spatial/navigational processing or, alternatively. It may also be the case that threatening scenes evoke fewer spatial contextual associations than neutral scenes, because the processing focuses more on the affective gist. A further important observation in PPA is that neutral bodies elicit higher activity than fearful bodies ($P < 0.0269$) compatible with body to scene affective coloring. The results in RSC show a similar pattern, but only when the bodies are presented in neutral scenes: neutral bodies in neutral scenes trigger more activity than fearful bodies in neutral scenes ($P < 0.0274$). There was no difference in RSC between fearful and neutral bodies when they were presented in a fearful or scrambled scene ($P < 0.2955$).

Finally, an area in the inferior occipital gyrus (IOG), in the region corresponding to the object selective lateral occipital complex (LOC (Malach et al., 1995)) was sensitive to body x scene interactions. Threatening or neutral scenes containing a fearful or neutral body elicit more activation compared to when the same scenes contained a control stimulus (triangle). The finding that the presence of a body in a natural scene boosts activation in IOG may reflect attentional processing consistent with recent findings from Peelen and colleagues showing that attention towards particular object classes like bodies mediates activity in LOC when

natural scenes are presented (Peelen et al., 2009). We found evidence for a body emotion effect in IOG, as there was higher activation for fearful than for neutral bodies ($P < 0.0445$). Again, this may point to increased visual attention to emotional stimuli, consistent with a recent study showing that emotional hand gestures trigger higher activity in LOC, compared to neutral ones (Flaisch, Schupp, Renner, & Junghofer, 2009).

In conclusion, our results show clear condition specific interactions between category selective scene and body areas and these interactions appear to be driven by the valence of the scenes and the body postures. This affective contamination between bodies and scenes observed here is compatible with early interaction between body and scene processing as well as with later convergence models. Previous reports that the emotional gist of a scene in which a face is presented influences the face specific N170 ERP component (Righart & de Gelder, 2006) suggest a rapid interaction between the two processing streams. Detailed information about the time course of this reciprocal contamination is needed to clarify this matter further.

References

- Bar, M. (2004). Visual objects in context. *Nat Rev Neurosci*, 5(8), 617-629.
- Bar, M., & Aminoff, E. (2003). Cortical analysis of visual context. *Neuron*, 38, 347-358.
- de Gelder, B., & Bertelson, P. (2003). Multisensory integration, perception and ecological validity. *Trends Cogn Sci*, 7(10), 460-467.
- de Gelder, B., & Van den Stock, J. (2008). Bodily Expressive Action Stimulus Test (BEAST).
- de Gelder, B., Van den Stock, J., Meeren, H. K., Sinke, C. B., Kret, M. E., & Tamietto, M. (2009). Standing up for the body. Recent progress in uncovering the networks involved in processing bodies and bodily expressions. *Neuroscience and Biobehavioral Reviews*.
- Downing, P. E., Jiang, Y., Shuman, M., & Kanwisher, N. (2001). A cortical area selective for visual processing of the human body. *Science*, 293(5539), 2470-2473.
- Epstein, R., & Kanwisher, N. (1998). A cortical representation of the local visual environment. *Nature*, 392(6676), 598-601.
- Epstein, R. A. (2008). Parahippocampal and retrosplenial contributions to human spatial navigation. *Trends Cogn Sci*, 12(10), 388-396.
- Flaisch, T., Schupp, H. T., Renner, B., & Junghofer, M. (2009). Neural systems of visual attention responding to emotional gestures. *Neuroimage*, 45(4), 1339-1346.
- Forman, S. D., Cohen, J. D., Fitzgerald, M., Eddy, W. F., Mintun, M. A., & Noll, D. C. (1995). Improved assessment of significant activation in functional magnetic resonance imaging (fMRI): use of a cluster-size threshold. *Magn Reson Med*, 33(5), 636-647.
- Goebel, R., Esposito, F., & Formisano, E. (2006). Analysis of functional image analysis contest (FIAC) data with brainvoyager QX: From single-subject to cortically aligned group general linear model analysis and self-organizing group independent component analysis. *Hum Brain Mapp*, 27(5), 392-401.
- Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: a module in human extrastriate cortex specialized for face perception. *J Neurosci*, 17(11), 4302-4311.
- Kret, M., & de Gelder, B. (submitted). Perceiving bodies in a social context.
- Lundqvist, D., Flykt, A., & Öhman, A. (1998). *The Karolinska Directed Emotional Faces - KDEF*. Stockholm: Karolinska Institutet.

- Malach, R., Reppas, J. B., Benson, R. R., Kwong, K. K., Jiang, H., Kennedy, W. A., et al. (1995). Object-related activity revealed by functional magnetic resonance imaging in human occipital cortex. *Proc Natl Acad Sci U S A*, 92(18), 8135-8139.
- Meeren, H. K., van Heijnsbergen, C. C., & de Gelder, B. (2005). Rapid perceptual integration of facial expression and emotional body language. *Proc Natl Acad Sci U S A*, 102(45), 16518-16523.
- Peelen, M. V., & Downing, P. E. (2005). Selectivity for the human body in the fusiform gyrus. *J Neurophysiol*, 93(1), 603-608.
- Peelen, M. V., Fei-Fei, L., & Kastner, S. (2009). Neural mechanisms of rapid natural scene categorization in human visual cortex. *Nature*, 460(7251), 94-97.
- Righart, R., & de Gelder, B. (2006). Context influences early perceptual analysis of faces--an electrophysiological study. *Cereb Cortex*, 16(9), 1249-1257.
- Talairach, J., & Tournoux, P. (1988). *Co-Planar Stereotaxic Atlas of the Human Brain*. New York: Thieme Medical Publishers.
- Tamietto, M., Castelli, L., Vighetti, S., Perozzo, P., Geminiani, G., Weiskrantz, L., et al. (2009). Unseen facial and bodily expressions trigger fast emotional reactions. *Proceedings of the National Academy of Sciences of the USA*, 106(42), 17661-17666.
- Thorpe, S., Fize, D., & Marlot, C. (1996). Speed of processing in the human visual system. *Nature*, 381(6582), 520-522.
- Van den Stock, J., Righart, R., & de Gelder, B. (2007). Body expressions influence recognition of emotions in the face and voice. *Emotion*, 7(3), 487-494.

Chapter 4: Body expressions influence recognition of facial expressions and emotional prosody⁴

Introduction

Imagine yourself taking a walk in the woods with a friend. You're talking to each other and suddenly he sees a snake approaching. His behavior changes immediately: his eyes open widely, his shoulders move backward, and his tone of voice changes. Over the last decade considerable progress has been made in understanding the functional and neuro-anatomical basis of human emotions. The combined findings from psychophysical investigations, lesion studies and brain imaging in neurologically intact observers have already provided a wealth of insights in how viewers process emotional information. Yet, it is noteworthy that our present understanding of how emotions are processed in humans is almost entirely based on studies investigating the perception of facial expressions (Adolphs, 2002).

Considering the emotional value of bodily expressions, it is somewhat surprising that the study of perception of whole body expressions lags so far behind that of facial expressions. Whole body expressions provide information about the emotional state of the producer, but also signal his action intentions. For example, a fearful body expression can signal the presence of a threat, but also how the producer intends to deal with it: flee, fight or freeze. Therefore, body expressions reveal a close link between emotion and (adaptive) behavior. Despite the early work of (Darwin) (1872), who described in detail the body expressions of many different emotions, there have been only a few isolated studies on human body postures in the past decades (Argyle, 1988; Ekman, 1965; Sprengelmeyer et al., 1999).

Our ability to perceive these emotional behaviors and how they are represented in the brain are now becoming important research topics. The findings so far have revealed striking similarities between how we process facial and whole body emotions (de Gelder, 2006).

For instance, at the behavioral level, some of the well-explored perceptual mechanisms involved in face processing also play a role in perception of bodies. Faces and bodies seem to be processed as invariant configurations inducing so called configural

⁴ This chapter is based on: Van den Stock, J., Righart, R., & de Gelder, B. (2007). Body expressions influence recognition of emotions in the face and voice. *Emotion*, 7(3), 487-494.

processing strategies, whereas other complex stimuli are more processed as an assemblage of features. Configural processing is often measured by the inversion effect (configural stimuli presented upside-down are more difficult to recognize than other complex inverted stimuli) and this effect has recently also been reported for bodies (Reed, Stone, Bozova, & Tanaka, 2003). Similar to the face inversion effect, the body inversion effect has also been measured with Event Related Potentials (ERP). A strong ERP inversion effect similar to what obtains for faces was observed for bodies (Stekelenburg & de Gelder, 2004).

Recent developmental findings now also underscore the important role of perceiving bodies for infants. For example, preferential processing of either faces or bodies might be a function of the distance to the stimulus. If a face is present at close range, especially the eyes are important, but when the distance increases, the configural properties of the whole face play a role (Johnson, 2005). This argument can be extended to whole bodies and suggests that whole body expressions are preferentially processed when the perceiver is further away from the stimulus. In line with this, behavioral data indicate specific expectations about the canonical properties of static faces and bodies at around 18 months (Slaughter, Stone, & Reed, 2004) and of dynamic bodies at three months (Bertenthal, Proffitt, & Kramer, 1987). Furthermore, ERP recordings provide evidence for similar processing of the configuration of faces and bodies at 3 months of age (Gliga & Dehaene-Lambertz, 2005).

At the functional neuro-anatomical level, a brain area in lateral occipital cortex has been described as responding selectively to neutral bodies or body parts (Downing, Jiang, Shuman, & Kanwisher, 2001). The fact that this area is very close to the motion sensitive area MT may explain its sensitivity to movement (Astafiev, Stanley, Shulman, & Corbetta, 2004), but see, (Peelen & Downing, 2005; Schwarzlose, Baker, & Kanwisher, 2005). Also, recent observations indicate significant proximity between faces and bodies in fusiform cortex (Schwarzlose, Baker, & Kanwisher, 2005) consistent with the finding that fearful bodies activate the face area in middle fusiform cortex (de Gelder, Snyder, Greve, Gerard, & Hadjikhani, 2004; Hadjikhani & de Gelder, 2003) and the finding that watching video images of angry hands and angry faces activate largely overlapping brain areas (Grosbras & Paus, 2006).

Recently, we proposed a model for the underlying circuitry of perception of emotional body language (de Gelder, 2006), advocating a two-system network, with the amygdala at the core: a primary subcortical one (including amygdala, striatum, pulvinar and superior colliculus) involved in rapid automated reflex-like perception of whole body expressions and a more cortically based one involved in explicit recognition. The latter

system comprises amygdala, superior temporal sulcus, presupplementary motor area, inferior parietal lobule and inferior frontal gyrus (Grèzes, Pichon, & de Gelder, 2006; Grosbras & Paus, 2006)

The present study investigates emotional body postures, how they are perceived and what their influence is on recognition of facial and vocal expressions of emotion. In Experiment 1 we investigate how well emotions are recognized from bodily expressions. In Experiment 2 we address the issue of synergies between facial expressions and bodily expressions. In Experiment 3 we explore the impact of bodily expressions on recognition of emotional voices.

Experiment 1: recognition of bodily expressions

The goal of this experiment was to test recognition of body expressions with a newly developed set of emotional body images. We asked participants to match a validated set of whole body expressions in a two alternative forced choice task. We used a matching task instead of a naming or categorization task because we wanted to investigate how well the different emotions are recognized on the basis of similarities with other stimuli in the same category and not mediated by the use of verbal labels.

Method

Participants. 17 neurologically intact volunteers, between the age of 18 and 28 (mean age was 21.3), participated in the experiment.

Materials and procedure. Materials consisted of 72 grayscale photographs representing semiprofessional actors (half male) expressing different emotions with their whole body (anger, fear, happiness and sadness) but with the face blurred. Selection of materials for use in the present experiment was based on the results of a pilot study in which the images were presented one by one on a screen and shown for 4000 ms with a 4000 ms interval. Participants were instructed to categorize each stimulus in a forced-choice procedure choosing one among five emotions as quickly and as accurately as possible and indicating the response on an answering sheet. For use in the present study, we only used images recognized above 70% accuracy.

A stimulus consisted of a target picture presented at the top and two probes left and right underneath (see Figure 1A for an example). There were always three different identities all three of the same gender and one of the probes had the same expression as the target. We

balanced the design in the way that for example when fear was the target expression, there were two trials (one with male actors and one with female actors) with an angry distracter, two trials with a happy distracter and two trials with a sad distracter. A total of 72 images was used, arranged in 24 trials (4 emotion categories x 3 distracter categories x 2 different genders). To avoid identity based matching, we used three different identities on each trial.

Stimuli were presented on a computer screen and participants were requested to match (as accurately and fast as possible) one of the bottom pictures to the one on top, based on similarity of expressed emotion. No instructions were given about which emotions could be expected. They responded by pressing the corresponding button, indicating their choice for the left or right probe. The stimulus was presented until response. During the 1000 ms intertrial interval, a blank screen was shown.

Results

Mean accuracy data are shown in Figure 1B, mean reaction time data in Figure 1C. One-sample *t*-tests show that recognition of all body emotions is above chance level (50%), $t(16) \geq 12.33, p \leq 0.001$.

A repeated measures ANOVA was carried out with expression (4 levels: anger, fear, happiness and sadness) as within subjects variable. This resulted in a significant effect, $F(3, 48) = 10.37, p < 0.001$. Bonferroni corrected post-hoc paired samples *t*-tests showed significant differences between anger and sadness, $t(16) = 3.79, p < 0.002$; fear and happiness, $t(16) = 4.40, p < 0.001$; and fear and sadness, $t(16) = 5.22, p < 0.001$.

A repeated measures ANOVA on the same data, but with expression of the distracter as within subjects factor also showed a significant effect, $F(3, 48) = 4.69, p < 0.006$. Bonferroni corrected post-hoc paired samples *t*-tests showed significant differences between anger and sadness as distracters, $t(16) = 3.85, p < 0.001$; and fear and sadness as distracters, $t(16) = 3.05, p < 0.008$. This is in line with the findings of the analysis with target body expression as within subjects variable, since it shows that angry and fearful bodily expressions are recognized less accurately than sad bodily expressions.

In order to find out which emotion the expression fear was most often confused with, we calculated the number of errors as a function of distracter emotion on the trials where fear was the target emotion. Seventy-nine percent of the errors were made when anger was the distracter, indicating fear was most frequently confused with anger.

We calculated the median reaction times per participant per condition and conducted a repeated measures ANOVA with expression (4 levels) as within subjects variable. This

showed a significant effect, $F(3, 48) = 17.18, p < 0.001$. Bonferroni corrected post hoc paired sampled t -tests revealed significant differences between anger and sadness, $t(16) = 5.10, p < 0.001$; fear and happiness, $t(16) = 3.92, p < 0.001$; and fear and sadness, $t(16) = 3.94, p < 0.001$.

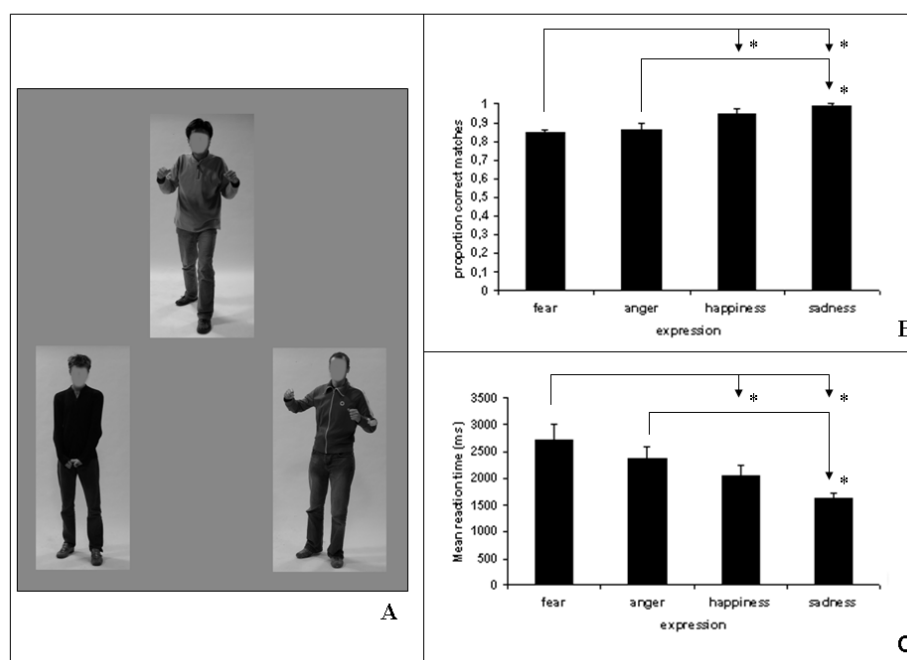


Figure 1. (A) Stimulus example of a trial in Experiment 1, showing an angry expression on top (target) and bottom right and a sad expression on the bottom left. Graphs show accuracy (B) and mean reaction times (C) as a function of expression. Error bars represent 1 SEM. $*p < 0.005$.

Discussion

The results of Experiment 1 indicate that the stimuli from our newly developed set body expressions are well recognizable without the help of verbal labels. The data also provide evidence for fear as the most difficult bodily expression to recognize in a forced choice paradigm. This finding has also been reported for facial expressions (Milders, Crawford, Lamb, & Simpson, 2003). Fearful expressions can be variable, depending on the kind of threat: one can be afraid of the dark, of getting hit, of making a public appearance, of being

rejected, etc. These different kinds of fear are associated with different defensive behaviors. This may explain why fearful whole body expressions are more difficult to recognize. However, Atkinson, Dittrich, Gemmell and Yong (2004) presented static and dynamic whole body expressions (face-blurred) at three levels of intensity and both in full light and point light displays. They asked participants to verbally label the stimuli in a five alternative forced choice task (anger, disgust, fear, happiness and sadness). For the static full light displays, they found anger to be more poorly recognized than fear, happiness and sadness, with little difference between the latter. In the present study, we found no significant difference between angry and fearful bodies, but fear was more poorly recognized than happiness and sadness (as indicated by both accuracy and reaction time data). Apart from the methodological differences (like for example the number of presented emotions and the type of task), the differences between emotions reported by Atkinson et al. may be due to differences in how well the stimuli are recognized. We accounted for this possibility by selecting the photographs on the basis of the results of a pilot study, in which we tested how well they are recognized.

Experiment 2: the influence of body expressions on recognition of facial expressions

Only one study has investigated the combined perception of human facial and bodily expressions (Meeren, van Heijnsbergen, & de Gelder, 2005). Participants were presented compound images of faces on bodies and their emotional content was either congruent or incongruent. The participants' task was to categorize the facial expression. Electrical brain responses were measured with EEG. The behavioral results showed participants were more accurate and faster when face and body expressed the same emotion. The ERP data provided evidence for an early perceptual integration of emotions expressed by face and body (around 115 ms post stimulus onset). Here, we extend on those earlier findings by investigating whether the effects observed by Meeren et al. (2005) also apply to the emotions fear and happiness (as opposed to fear and anger). Further, morphed faces were used in order to test whether individuals use information from bodies differently when facial expressions are ambiguously positioned in-between fear and happy.

Method

Participants. Participants were 14 first year psychology students (mean age 19.1 years).

Materials and procedure. Grayscale photographs of a male actor with a fearful and happy body expression were selected from our own validated database (recognized correctly 100 and 90% respectively). One identity of the Ekman and Friesen (1976) facial expressions database was selected. We used the happy and fearful expression as extremes to create a 5-step continuum between the two expressions. The morphing of the expressions was done according to the procedure developed by Benson and Perrett (1991). The faces were edited in size and pasted on the body to create a realistically looking “identity” (see Figure 2B, for examples). Every facial expression was paired with every bodily expression. This resulted in 10 compound stimuli: the five facial expressions pasted on the happy bodily expression and the same 5 faces pasted on the fearful bodily expression.

All compound stimuli were presented on a computer screen 15 times in random order in 3 identical blocks. Presentation time was 150 ms, after which a blank screen appeared.

Participants were instructed to indicate whether the face expressed fear or happiness. Intertrial interval was 2000 ms.

Results

The proportion *happy* responses was calculated for each participant and for each compound stimulus. Results are displayed in Figure 2A. A 5 x 2 repeated measures ANOVA was carried out with face (5 levels) and body (2 levels) as within subjects variables. This revealed a main effect of face, $F(4, 52) = 106.65, p < 0.001$, body, $F(1, 13) = 37.56, p < 0.001$ and a significant interaction, $F(4, 52) = 4.78, p < 0.002$. To follow up on the interaction effect we compared for each of the 5 facial expressions, the difference between the proportions ‘happy’ responses as a function of the accompanying bodily expression. Bonferroni corrected t tests showed a significant difference on three adjacent levels of the facial expression continuum, starting from the fear end ($p < 0.006$; $p < 0.001$; $p < 0.001$ respectively). The other two differences were not significant ($p < 0.018$ and $p < 0.265$).

A trend analysis showed there was a linear trend in the face factor, $F(1,13) = 554.33, p < 0.001$, indicating the distances between the face morphs were perceived as equal. The body x

face interaction showed a quadratic trend, $F(1,13) = 23.65$, $p < 0.001$, indicating the influence of the body is smaller at the extreme ends of the face continuum.

We calculated the median reaction times by participant and condition and performed the same 2×5 ANOVA on the reaction time data. This revealed no significant effects, indicating the results are not biased by a speed-accuracy tradeoff.

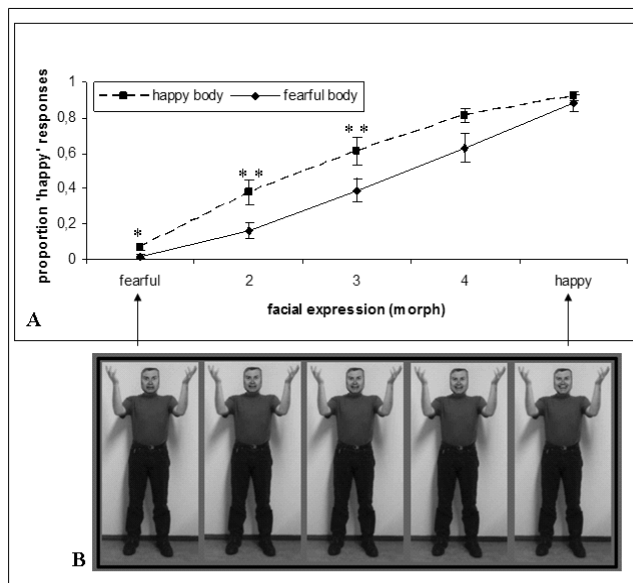


Figure 2. (A) Mean proportion happy responses in Experiment 2 as a function of facial expression. Error bars represent 1 standard error around the mean. * $p < 0.01$; ** $p < 0.001$. (B) Stimulus examples of Experiment 2, showing a happy body expression with a morphed face (left: 100% fearful; right 100% happy) .

Discussion

The results of Experiment 2 provide clear evidence that recognition of facial expressions is influenced by the accompanying body language. A happy face on a happy body is categorized more frequently as happy, compared to when the same happy face appears on a fearful body. And a fearful face on a fearful body is categorized as more fearful, compared to when it appears in combination with a happy body expression. It should be stressed that the instructions explicitly stated to categorize the facial expression, so there was no ambiguity as regarded to the 'target' for classification.

These results are consistent with a previous study using compound stimuli of angry and fearful facial and whole body expressions (Meeren et al., 2005) and extend those findings to the emotions fear and happiness.

Moreover, the interaction and trend analysis reported in this study indicate that the influence of the body expression is a function of the ambiguity of the facial expression: the whole body expression has the most influence when the face ambiguity is highest and decreases with reduced facial ambiguity.

Experiment 3: The influence of body language on recognition of voice prosody

Multisensory integration is considered adaptive, since it reduces stimulus ambiguity (de Gelder & Bertelson, 2003). Previous studies have indicated that facial expressions and emotional tone of voice or emotional prosody influence each other (de Gelder & Vroomen, 2000; Massaro & Egan, 1996). Emotional prosody refers to the variations in melody, intonations, pauses, stresses and accents of speech. Factors that play a role in voice prosody of some emotions are duration and intonation. For example a happy sentence is of normal duration, the pitch is high and there is a major change in pitch. Acoustically, 'fear' is very similar to happiness and the duration is also normal, the mean pitch is also high but the change in pitch is smaller than in a happy sentence. Integration of affective information from different sensory channels seems to be essential for accurate and fast recognition of emotions.

Developmental studies on recognition of prosody typically present facial expressions with either prosodic congruent or incongruent vocal expressions, while measuring the looking time at the faces (Soken & Pick, 1992; Walker-Andrews, 1986; A.S. Walker, 1982; A. S. Walker & Grolnick, 1983). The results indicate that already at three months of age, infants can detect changes in prosody (A. S. Walker & Grolnick, 1983).

In a study with static facial expressions and emotional spoken sentences, de Gelder and Vroomen (2000) observed a crossmodal influence of the affective information. Recognition of morphed vocal expressions was biased towards the simultaneously presented facial expression, even when the participants were instructed to ignore the visual stimuli. A follow up study suggests that this crossmodal integration of affective information takes place automatically, independent of attentional factors (Vroomen, Driver, & de Gelder, 2001) and

works also when the observer is unaware of the expression of the face, as observed in a cortically blind patient (de Gelder, Morris, & Dolan, 2005). Investigations of the time course of this integration with ERP have indicated that affective information from different sensory channels is combined early in the perceptual process (de Gelder, Bocker, Tuomainen, Hensen, & Vroomen, 1999). In Experiment 3, we used a similar paradigm as de Gelder and Vroomen (2000) (experiment 3), but we tested for the effect of whole body expressions instead of facial expressions.

Method

Participants. The group consisted of 16 neurologically intact participants (mean age 32.9).

Materials and procedure. The visual stimuli were the same fearful and happy whole body expressions as in experiment 2, but with the faces blurred (see Figure 3A for an example). The auditory stimulus materials consist of a spoken sentence, edited as to express different levels of emotion on a 7-step continuum between fearful and happy. The editing consisted of adjusting of the duration, pitch range and pitch register (see (de Gelder & Vroomen, 2000), for details). For the present study, we only used the last 4 words of the sentence (“kwam met het vliegtuig”, meaning “arrived by plane”). The auditory stimuli lasted about 600ms.

Audiovisual stimuli were created by pairing each body expression with each of the seven auditory stimuli, thus resulting in 14 audiovisual stimuli: the 7 vocal expressions paired with the fearful body and the same 7 vocal expressions paired with the happy body.

The visual stimuli were presented on a computer screen. Auditory stimuli were presented at a comfortable listening level over loudspeakers. The on- and offset of the visual stimulus was synchronized with the auditory stimulus. Participants had a maximum of 4000ms to respond, followed by an intertrial interval of 1000ms.

The task was to categorize the expression of the voice (fearful or happy) in a two alternative forced choice task. On catch trials, a white “X” appeared on the body and participants were to refrain from responding. We included these catch trials to make sure participants saw the body. The experiment was run in two sessions of both 49 randomized trials (3 presentations of all 14 audiovisual stimuli + 7 catch trials). The sessions were preceded by 10 practice trials.

Results

Only the participants that missed no more than 5 catch trials were selected for the analyses. For this reason, 2 participants were excluded from the analysis. The proportion of *happy* responses was calculated for each participant for every combination of voice prosody and body expression on the experimental trials. Results are displayed in Figure 3B.

The data were submitted to a repeated measures ANOVA with voice (7 levels) and body (2 levels) as within subject factors. This showed a main effect of voice, $F(6, 78) = 24.90, p < 0.001$, and body, $F(1, 13) = 9.94, p < 0.008$, but no interaction.

We calculated the median reaction times by participant and condition. Reaction time data are shown in Figure 3C. The 7 x 2 repeated measures ANOVA on the reaction time data only showed a main effect of voice, $F(6, 78) = 9.43, p < 0.001$. This simply reflects the fact that reaction times increase as the vocal expression becomes more ambiguous, as can be seen in Figure 3C.

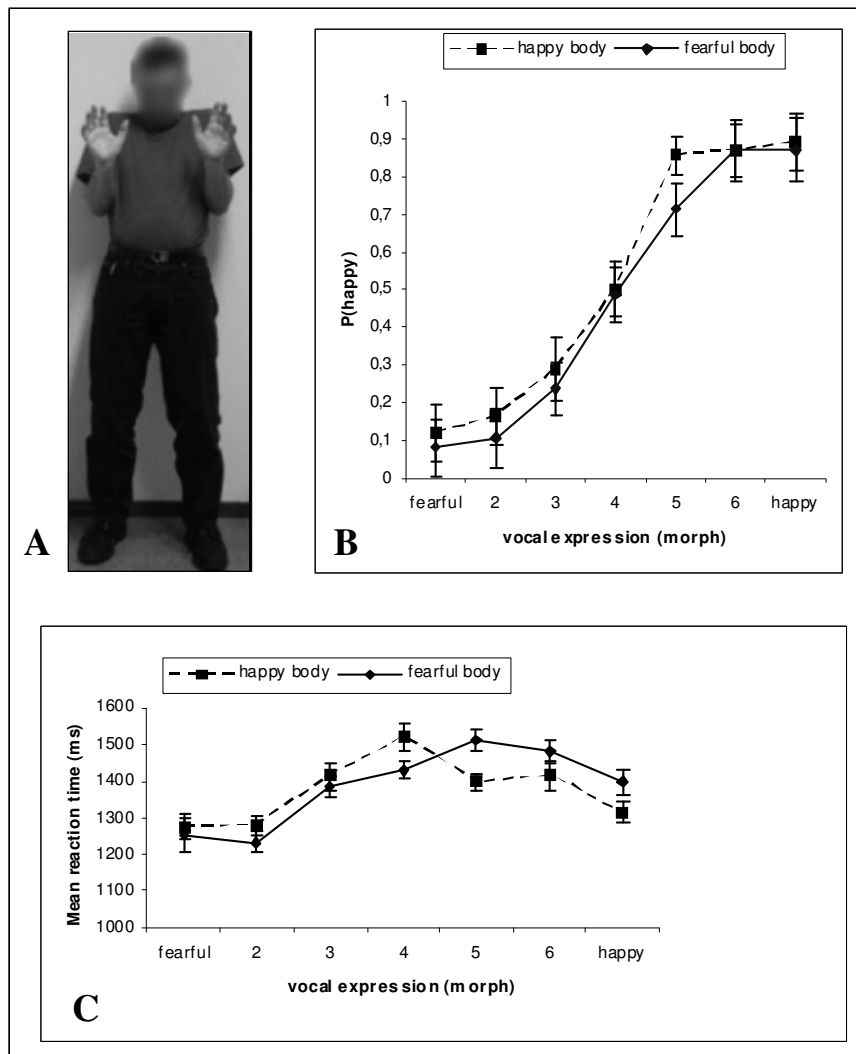


Figure 3. (A) Fearful whole body expression as presented in Experiment 3. (B) Mean proportion happy responses in Experiment 3 as a function of vocal expression. (C) Mean reaction time in Experiment 3 as a function of vocal expression. Error bars represent 1 standard error around the mean.

Discussion

The results of experiment 3 indicate an influence of a perceived whole body expression on the recognition of voice prosody. When observers make judgments about the

emotion conveyed in a voice, recognition is biased towards the simultaneously perceived whole body expression. The task required attention to be focused towards the voice, but nevertheless, there is a systematic influence of body expression.

The crossmodal affective bias effect has also been observed between voice prosody and facial expressions and seems to be mandatory and automatic (de Gelder & Vroomen, 2000; Vroomen et al., 2001). Vroomen et al. (2001) investigated whether bimodal integration of affective faces and voices required limited attentional resources. Subjects judged whether a voice expressed happiness or fear, while instructed to ignore a concurrently presented static facial expression. Additional tasks were added, to manipulate the attentional load. However, the crossmodal bias effect was independent of whether the subjects performed a demanding attentional task. In line with this, the present experiment suggests perceptual integration of bimodal emotion expression rather than integration of the two sources based on a later post-perceptual and more cognitive process as suggested by previous literature. ERP-studies with audiovisual affective stimuli point to an early integration of sensory modalities (within 110 ms post-stimulus onset), also suggesting a perceptual mechanism, instead of a later more cognitive process (Pourtois, de Gelder, Vroomen, Rossion, & Crommelinck, 2000). A study with intracranial recordings in monkeys indicated integration of facial and vocal signals in primary auditory cortex through enhancement and suppression of field potentials (Ghazanfar, Maier, Hoffman, & Logothetis, 2005). Most importantly, the combination of a fearful face with a fearful tone of voice increases activation in amygdala (Dolan, Morris, & de Gelder, 2001) indicating that the merging of information across stimulus categories is driven by the perception of the meaning irrespective of the medium through which the meaning is conveyed. These questions need to be addressed in follow-up studies using methods that provide a better insight in the temporal dynamics.

General discussion

Experiment 1 showed that the newly developed stimuli are easily recognized when no verbal labels are provided. We also found fear to be the most difficult emotion to be recognized, consistent with previous reports on facial expressions (Milders et al., 2003).

In Experiment 2, we found perception of facial expressions to be biased towards the not explicitly attended and task irrelevant body language. This replicates our findings of a similar previous study (Meeren et al., 2005), but additionally extends the observations to other

combinations of emotions and indicates that the magnitude of the influence of the body expression depends on the facial expression ambiguity. The data of Experiment 3 show that when participants are asked to identify the emotional tone of a voice, while ignoring a simultaneously presented body, they are nevertheless susceptible to be influenced by the bodily expression.

The results of the present study clearly indicate the importance of whole body expressions as significant emotional stimuli and reveals similarities with findings from facial expression research. The presence of an unattended expressive body influences recognition of faces and auditory stimuli.

From an evolutionary perspective, an important adaptive function of body language is communication of relevant information to other members of the species. Especially in social species, there are considerable adaptive benefits in the ability to interpret emotional displays by conspecifics (Dawkins & Krebs, 1978). This is especially the case when the facial expression of the producer is not visible, for example because of the viewers' perspective or because of a too great distance to the stimulus.

In the case of emotional body language, perceiving dynamics seem to be particularly important. Recognition of dynamic whole body expressions is easier than static stimuli (Atkinson et al., 2004) and seems little affected by cultural factors (Hejmadi, Davidson, & Rozin, 2000; Rozin, Taylor, Ross, Bennett, & Hejmadi, 2005). The present studies used static images, in line with the large majority of studies of facial expressions, but there is reason to believe that the important dynamic emotion information may not need to be present explicitly to create a dynamic percept. When viewing two successive presentations of a stimulus object with implied motion, subjects fail to notice the difference between them if the second one represents the same event, but a moment later in time (Freyd, 1983). Moreover, viewing implied motion stimuli activates brain area MT/MST, involved in the processing of movement (Kourtzi & Kanwisher, 2000). For the case of emotional whole body expressions, we observed that viewing static fear images yield strong activity in motor areas (de Gelder et al., 2004). More recently, we compared activation for static versus dynamic presentation of the same images and observed no difference in amygdala activity for the two presentation conditions (Grèzes et al., 2006). Thus there is reason to believe that the sight of a bodily expression of emotion affects the viewer profoundly even when motion is not explicitly shown. Creating this emotional movement illusion is indeed what the visual arts have excelled at since a very long time.

References

- Adolphs, R. (2002). Neural systems for recognizing emotion. *Curr Opin Neurobiol*, 12(2), 169-177.
- Argyle, M. (1988). *Bodily communication*. London: Methuen.
- Astafiev, S. V., Stanley, C. M., Shulman, G. L., & Corbetta, M. (2004). Extrastriate body area in human occipital cortex responds to the performance of motor actions. *Nat Neurosci*, 7(5), 542-548.
- Atkinson, A. P., Dittrich, W. H., Gemmell, A. J., & Young, A. W. (2004). Emotion perception from dynamic and static body expressions in point-light and full-light displays. *Perception*, 33(6), 717-746.
- Benson, P. J., & Perrett, D. (1991). Perception and recognition of photographic quality facial caricatures: Implications for the recognition of natural images. *Eur J Cog Psychology*, 3, 105-135.
- Bertenthal, B. I., Proffitt, D. R., & Kramer, S. J. (1987). Perception of biomechanical motions by infants: implementation of various processing constraints. *J Exp Psychol Hum Percept Perform*, 13(4), 577-585.
- Darwin, C. (1872). *The expression of the emotions in man and animals*. London: John Murray.
- Dawkins, R., & Krebs, J. R. (1978). Animals' signals: information or manipulation. In J. R. Krebs & N. B. Davies (Eds.), *Behavioural Ecology: An Evolutionary Approach*. (pp. 282-309). Oxford: Blackwell Scientific Publications.
- de Gelder, B. (2006). Towards the neurobiology of emotional body language. *Nature Reviews Neuroscience*, 7(3), 242-249.
- de Gelder, B., & Bertelson, P. (2003). Multisensory integration, perception and ecological validity. *Trends Cogn Sci*, 7(10), 460-467.
- de Gelder, B., Bocker, K. B., Tuomainen, J., Hensen, M., & Vroomen, J. (1999). The combined perception of emotion from voice and face: early interaction revealed by human electric brain responses. *Neurosci Lett*, 260(2), 133-136.

- de Gelder, B., Morris, J. S., & Dolan, R. J. (2005). Unconscious fear influences emotional awareness of faces and voices. *Proc Natl Acad Sci U S A*, 102(51), 18682-18687.
- de Gelder, B., Snyder, J., Greve, D., Gerard, G., & Hadjikhani, N. (2004). Fear fosters flight: A mechanism for fear contagion when perceiving emotion expressed by a whole body. *Proc Natl Acad Sci U S A*, 101(47), 16701-16706.
- de Gelder, B., & Vroomen, J. (2000). The perception of emotions by ear and by eye. *Cognition and Emotion*, 14(3), 289-311.
- Dolan, R. J., Morris, J. S., & de Gelder, B. (2001). Crossmodal binding of fear in voice and face. *Proc Natl Acad Sci U S A*, 98(17), 10006-10010.
- Downing, P. E., Jiang, Y., Shuman, M., & Kanwisher, N. (2001). A cortical area selective for visual processing of the human body. *Science*, 293(5539), 2470-2473.
- Ekman, P. (1965). Differential communication of affect by head and body cues. *J Pers Soc Psychol*, 2(5), 726-735.
- Ekman, P., & Friesen, W. V. (Artist). (1976). *Pictures of facial affect*
- Freyd, J. J. (1983). The mental representation of movement when static stimuli are viewed. *Percept Psychophys*, 33(6), 575-581.
- Ghazanfar, A. A., Maier, J. X., Hoffman, K. L., & Logothetis, N. K. (2005). Multisensory integration of dynamic faces and voices in rhesus monkey auditory cortex. *J Neurosci*, 25(20), 5004-5012.
- Gliga, T., & Dehaene-Lambertz, G. (2005). Structural encoding of body and face in human infants and adults. *Journal of Cognitive Neuroscience*, 17, 1328-1340.
- Grèzes, J., Pichon, S., & de Gelder, B. (2006). Perceiving fear in dynamic body expressions. *Neuroimage*, in press.
- Grosbras, M. H., & Paus, T. (2006). Brain networks involved in viewing angry hands or faces. *Cereb Cortex*, 16(8), 1087-1096.
- Hadjikhani, N., & de Gelder, B. (2003). Seeing fearful body expressions activates the fusiform cortex and amygdala. *Curr Biol*, 13(24), 2201-2205.

- Hejmadi, A., Davidson, R. J., & Rozin, P. (2000). Exploring Hindu Indian emotion expressions: evidence for accurate recognition by Americans and Indians. *Psychol Sci*, 11(3), 183-187.
- Johnson, M. H. (2005). Subcortical face processing. *Nat Rev Neurosci*, 6(10), 766-774.
- Kourtzi, Z., & Kanwisher, N. (2000). Activation in human MT/MST by static images with implied motion. *J Cogn Neurosci*, 12(1), 48-55.
- Massaro, D. W., & Egan, P. B. (1996). Perceiving Affect from the Voice and the Face. *Psychonomic Bulletin and Review*, 3, 215-221.
- Meeren, H. K., van Heijnsbergen, C. C., & de Gelder, B. (2005). Rapid perceptual integration of facial expression and emotional body language. *Proc Natl Acad Sci U S A*, 102(45), 16518-16523.
- Milders, M., Crawford, J. R., Lamb, A., & Simpson, S. A. (2003). Differential deficits in expression recognition in gene-carriers and patients with Huntington's disease. *Neuropsychologia*, 41(11), 1484-1492.
- Peelen, M. V., & Downing, P. E. (2005). Selectivity for the human body in the fusiform gyrus. *J Neurophysiol*, 93(1), 603-608.
- Pourtois, G., de Gelder, B., Vroomen, J., Rossion, B., & Crommelinck, M. (2000). The time-course of intermodal binding between seeing and hearing affective information. *Neuroreport*, 11(6), 1329-1333.
- Reed, C. L., Stone, V. E., Bozova, S., & Tanaka, J. (2003). The body-inversion effect. *Psychol Sci*, 14(4), 302-308.
- Rozin, P., Taylor, C., Ross, L., Bennett, G., & Hejmadi, A. (2005). General and specific abilities to recognise negative emotions, especially disgust, as portrayed in the face and the body. *Cognition and Emotion*, 19(3), 397-412.
- Schwarzlose, R. F., Baker, C. I., & Kanwisher, N. (2005). Separate face and body selectivity on the fusiform gyrus. *J Neurosci*, 25(47), 11055-11059.
- Slaughter, V., Stone, V. E., & Reed, C. (2004). Perception of Faces and Bodies: Similar or Different? *Current Directions in Psychological Science*, 13(6), 219-223.

- Soken, N. H., & Pick, A. D. (1992). Intermodal perception of happy and angry expressive behaviors by seven-month-old infants. *Child Development*, 63(787-795).
- Sprengelmeyer, R., Young, A. W., Schroeder, U., Grossenbacher, P. G., Federlein, J., Buttner, T., et al. (1999). Knowing no fear. *Proc Biol Sci*, 266(1437), 2451-2456.
- Stekelenburg, J. J., & de Gelder, B. (2004). The neural correlates of perceiving human bodies: an ERP study on the body-inversion effect. *Neuroreport*, 15(5), 777-780.
- Vroomen, J., Driver, J., & de Gelder, B. (2001). Is cross-modal integration of emotional expressions independent of attentional resources? *Cogn Affect Behav Neurosci*, 1(4), 382-387.
- Walker-Andrews, A. S. (1986). Intermodal perception of expressive behaviors: relation of eye and voice? *Developmental Psychology*, 22, 373-377.
- Walker, A. S. (1982). Intermodal perception of expressive behaviors by human infants. *Journal of Experimental Child Psychology*, 33, 514-535.
- Walker, A. S., & Grolnick, W. (1983). Discrimination of vocal expressions by young infants. *Infant Behavior and Development*(6), 491-498.

Chapter 5: Human and animal sounds influence recognition of body language⁵

Introduction

When Hitchcock shows Norman Bates stabbing his victim to death in the shower or when the dorsal fin of a shark surfaces in “Jaws”, the soundtrack is always there to underscore the message. Movie directors rely heavily on the extra dimension added to the movie experience by the soundtrack to convey emotion and aim at creating a multimodal experience in the viewer.

Experimental research on combined perception of auditory and visual stimuli has a long history (Müller, 1840), and there is now considerable evidence that multisensory stimuli presented in spatial or temporal proximity are bound by the brain into a unique perceptual gestalt (for reviews see de Gelder & Bertelson, 2003; Welch & Warren, 1986). Studies investigating the recognition of bimodal human emotional expressions typically consist of presenting audiovisual stimulus pairs in which the emotional content between the visual and auditory modality is either congruent or incongruent (de Gelder, Bocker, Tuomainen, Hensen, & Vroomen, 1999; de Gelder & Vroomen, 2000; Ethofer et al., 2006; Massaro & Egan, 1996; Spreckelmeyer, Kutas, Urbach, Altenmüller, & Munte, 2006; Van den Stock, Righart, & de Gelder, 2007). For example, de Gelder and Vroomen (2000) presented a static face expressing sadness or happiness combined with a spoken sentence with an emotionally neutral meaning but with either a sad or happy tone of voice. Participants were asked to ignore the voice and to indicate whether the face expressed happiness or sadness. The results indicated a clear crossmodal bias, e.g. a sad facial expression paired with a happy voice was recognized more as happy, compared to when the same facial expression was paired with a sad voice. In a follow up experiment, the task was reversed and participants were instructed to categorize the vocal expression and ignore the face. The results showed that the voice ratings were biased towards the emotion expressed by the face. The findings from de Gelder and Vroomen (2000) are consistent with other studies on bimodal perception of affect expressed in face and voice (de Gelder et al., 1999; Ethofer et al., 2006; Massaro & Egan, 1996).

⁵ This chapter is based on: Van den Stock, J., Grèzes, J., de Gelder, B. (2008). Human and animal sounds influence recognition of body language. *Brain Research*, 1242, 185-190.

We know from daily experience that emotions are not solely expressed in the face and the voice, but also conveyed very forcefully and over considerable distance by postures and movements of the whole body. Research on whole body perception is emerging as a new field in neuroscience (e.g. Atkinson, Dittrich, Gemmell, & Young, 2004; de Gelder, 2006; Grezes, Pichon, & de Gelder, 2007; Peelen & Downing, 2007). In view of these new findings a question is whether similar interactions as previously observed for facial expressions and auditory stimuli will also be obtained when observers are shown body-voice pairs. Recently, we presented static happy and fearful whole body expressions with faces blurred and each combined with a happy or fearful voice. Participants were asked to ignore the body expression and rate the emotion expressed by the voice. The results indicated that recognition of voice prosody was biased towards the emotion expressed by the whole body (Van den Stock et al., 2007, experiment 3). Here, we take that line of research a step further and investigate whether similar effects can be obtained with a dynamic body images. Also, we address the question whether as suggested by the familiar movie viewer's experience there is crossmodal influence if both modalities are unmistakably and recognizably produced by a different source as is indeed often the case in naturalistic circumstances in which we observe the bodily expressions with the audio track accompanying the visual images.

In this study, we present dynamic whole body expressions of emotion, showing persons engaged in an everyday activity and in a realistic context. In contrast to earlier studies we used nonverbal auditory information consisting of human vocalizations and also of animal sounds, two conditions that befit the naturalistic circumstances of viewing emotional body expressions from a relative distance. By using these two kinds of auditory information we may address the issue whether environmental sounds (i.e. auditory stimuli originating from a source other than the visual stimulus) have a similar influence on recognition of visual human expressions as we expect voices to have.

Thirdly, to minimize semantic or verbal processing, which is initiated automatically when verbal information is presented, we used nonverbal auditory materials. Until now, only verbal vocalizations have been used to investigate crossmodal bias effects in processing human expressions. Nonverbal utterances have been used recently in scene-voice pairs. Spreckelmeyer et al. (2006) presented an emotionally sung syllable ("ha") paired with an emotional scene and asked participants to rate the valence of the scene. The authors did not observe an influence of the nonverbal vocalization on the ratings of the visual stimulus. However, pairing scenes with a sung syllable has limited ecological value. Also, a number of

scenes in this study evoke an emotional experience, rather than showing an emotional expression (for example a picture of a baby or bunny).

Here, we investigate the influence of human and environmental emotional auditory information on the recognition of emotional body expression. For the case of the environmental auditory stimuli, we presented animal vocalizations inducing fear or happiness, creating realistic bimodal stimuli in the congruent conditions. Participants were presented video clips of happy or fearful body language. These were simultaneously presented with either congruent or incongruent human or animal vocalizations, or without auditory information. The experiment used a two alternative forced choice task and the instructions requested the participants to categorize the emotion expressed by the body stressing speed and accuracy.

Method

Participants. Twenty-seven adults (14 male; 23 right-handed; mean age 31.5, range 18-50) participated in the experiment. They all gave written consent according to the Declaration of Helsinki. None of them had a history of neurological or psychiatric disorders. All had normal or corrected to normal vision and normal hearing.

Stimulus materials.

Visual stimuli. Video recordings were made of 12 semi-professional actors (6 women), coached by a professional director. They were instructed to approach a table, pick up a glass, drink from it and to put it back on the table. They performed this action once in a happy and once in a fearful manner. A continuous fragment of 800 ms was selected from each video showing the actor grasping the glass. Facial expressions were blurred using motion tracking software. See Figure 1 for an example.

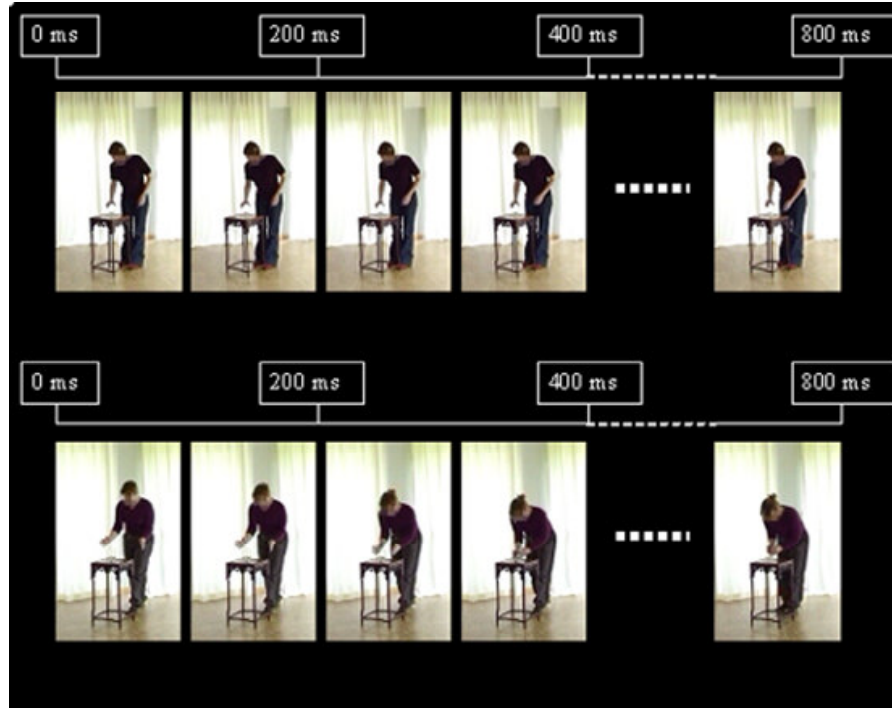


Figure 1. An example of frames from the video clips. The frame selection on the top row shows an actor grasping a glass in a fearful manner, the one on the bottom row performs the same action in a happy manner. The faces are blurred to minimize facial expression.

In a pilot study the 24 edited dynamic stimuli (20 frames) were presented 4 times to 14 participants. Participants were instructed to categorize as accurately and as fast as possible the emotion expressed by the actor (fear or happiness). The pilot session was preceded by eight familiarization trials. Sixteen stimuli were selected (2 gender x 4 actors x 2 emotions). Since we expected that recognition of the body language improves when the body stimuli are combined with congruent auditory information, body stimuli that were recognized at ceiling were not selected. Mean recognition of the selected stimuli was 86.1 percent (SD 9.7). A paired *t*-test between the fearful and happy body language showed no significant difference, $t(13) = 1.109, p < .287$.

Auditory stimuli. Audio recordings were made at a sampling rate of 44.1 kHz of 22 subjects (14 women), while they made non-verbal emotional vocalizations (fearful and happy). Specific scripts were provided for every target emotion. For example, for fear the

actors were instructed to imagine they were going to be attacked by a robber and to react to such an encounter in a nonverbal manner. Video clips were edited and the most representative 800 ms fragment from each recording was cut and digitally stored on a PC. In a pilot study the sounds were presented to 15 participants. Every sound was presented 4 times in a randomized order. The participants were instructed to categorize as accurately and as fast as possible the emotion expressed by the voice (fear or happiness). The pilot session was preceded by 3 familiarization trials. Based on these results, eight fearful and eight happy sounds were selected. Mean recognition of the stimuli was 94.6 percent (SD 6.7). A paired t -test between the fearful and happy vocalizations showed no significant difference, $t(14) = 0.474, p < .643$.

Environmental sounds consisted of aggressive dog barks and joyful bird songs and were downloaded from the internet. Stimuli were selected on the basis of their emotion inducing characteristics. In a third pilot study, these sounds were presented 4 times to 13 participants. They were instructed to categorize as accurately and as fast as possible the emotion induced by the sound (fear or happiness). The pilot session was preceded by 3 familiarization trials. Eight fear inducing and eight happiness inducing sounds were selected. Mean recognition of the stimuli was 94.8 percent (SD 5.7). A paired t -test between the fearful and happy vocalizations showed no significant difference, $t(12) = 1.469, p < .168$.

For each emotion we compared the ratings of the animal vocalizations with those of the human vocalizations. Independent samples t -tests showed no differences between the pairs $t(26) \leq 1.195, p < .243$.

Experimental stimuli were then constructed with these visual and auditory materials. For this purpose each video file was paired once with a fearful and happy human vocalization, resulting in a total of 32 bimodal stimuli (human video/human audio) and once with a fear (dog barking) and happiness inducing (birdsong) animal vocalization, resulting in a total of 32 bimodal stimuli (human video/animal audio).

Procedure. The experiment consisted of a visual (V) and an audio-visual (AV) block. In each block all stimuli were presented twice in random order. The order of the blocks was counterbalanced. The AV-block consisted of 128 trials (2 presentations of 64 stimuli: 16 fearful videos with congruent human sounds, 16 fearful videos with incongruent human sounds, 16 videos with congruent animal sounds and 16 videos with incongruent animal sounds), the V-block of 32 trials (2 presentations of 16 stimuli, 8 fearful and 8 happy clips). A

trial started with the presentation of a white fixation cross in the center of the screen against a dark background. The fixation cross had a variable duration to reduce temporal predictability (2000-3000 ms) and was followed by presentation of a stimulus (800 ms) after which a question mark appeared until the participant responded. A two alternative forced choice task was used requiring the participants to categorize the emotion expressed in the body by pressing the corresponding button (happy or fearful). Response buttons were counterbalanced across participants. Because we wanted to make sure participants saw the full length of the stimulus before they responded, they were instructed only to respond when the question mark appeared.

Results

Trials with reaction times below 1000 ms and above 3000 ms (post stimulus onset) were excluded. One participant responded outside this time window on more than 10 % of the trials and was therefore excluded from the analysis. We computed the proportion happy responses of the different conditions. Results are shown in Figure 2.

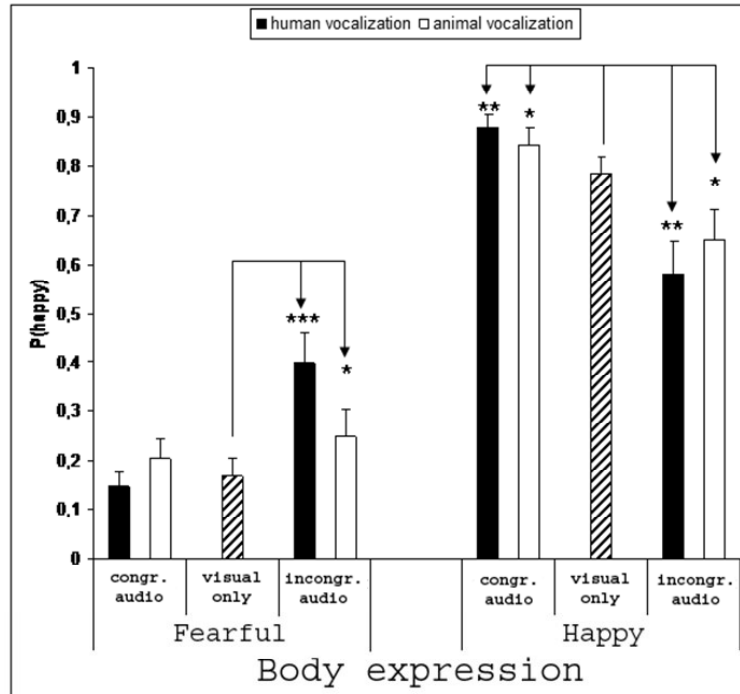


Figure 2. Proportion 'happy' in the bimodal and unimodal conditions, separated by emotion, auditory category and congruence. Error bars represent 1 SEM. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

Human vocalizations:

A repeated measures ANOVA was performed on the proportion happy responses with visual emotion (fearful and happy) and (human) auditory emotion (fearful, happy and no auditory stimulus) as within-subjects factors. This revealed a significant effect of visual emotion $F(1,25) = 85.993, p < .001$, auditory emotion, $F(2,50) = 16.523, p < .001$, and a significant interaction between visual emotion and auditory emotion, $F(2,50) = 5.761, p < .006$.

To follow up on the interaction effect and to test the influence of the auditory stimuli on the recognition of the visual stimuli, we performed paired sample *t*-tests. Against the background of our previous experiments using faces and voices (e.g. de Gelder & Vroomen, 2000), we expect expression recognition performance on the congruent stimulus combinations to be better, compared to the unimodal combinations. Likewise, performance on the unimodal conditions is expected to be higher than on the incongruent conditions. Therefore, we performed one-tailed *t*-tests, comparing the unimodal conditions (V) with their respective bimodal

conditions. For the happy body language, there was a difference between baseline [V(happy)] and both congruent AV, $t(25) = 2.935, p < .01$, and incongruent AV, $t(25) = 2.945, p < .01$. For the fearful body language, there was a significant difference between baseline [V(fear)] and incongruent AV, $t(25) = 4.217, p < .001$.

Animal vocalizations:

A repeated measures ANOVA on the proportion happy responses with visual emotion (fearful and happy) and (animal) auditory emotion (fearful, happy and no auditory stimulus) as within-subjects factors, revealed a significant effect of visual emotion $F(1,25) = 92.050, p < 0.001$, auditory emotion, $F(2,50) = 3.405, p < .041$, and an interaction between visual emotion and auditory emotion, $F(2,50) = 5.040, p < .010$. The post-hoc paired t -tests (one tailed) showed significant differences between V(happiness) and congruent AV, $t(25) = 1.823, p < 0.040$; between V(happiness) and incongruent AV, $t(25) = 1.948, p < 0.032$ and between V(fear) and incongruent AV, $t(25) = 1.726, p < 0.050$.

human and animal vocalizations:

To compare the influence of human with animal vocalizations, we ran a 2 (video emotion: fearful and happy) x 2 (auditory emotion: fearful and happy) x 2 (auditory source: human and animal) repeated measures ANOVA on the proportion happy responses. This revealed a significant main effect of visual emotion $F(1,25) = 56.048, p < 0.001$; auditory emotion $F(1,25) = 11.001, p < 0.005$, a two-way visual x auditory emotion interaction $F(1,25) = 11.564, p < 0.005$; a two-way auditory emotion x source interaction $F(1,25) = 16.088, p < 0.001$, a two-way visual emotion x source interaction $F(1,25) = 5.140, p < 0.05$; and a three-way visual emotion x auditory emotion x source interaction $F(1,25) = 5.532, p < 0.05$. The two-way auditory emotion x source interaction indicates a different influence of the human and animal vocalizations. To follow up on this effect, we compared the influence of the human with the animal vocalizations, by computing the difference between the congruent and incongruent combinations, for the human and animal sounds separately (namely the human congruent conditions minus the human incongruent conditions and the animal congruent conditions minus the animal incongruent conditions). This difference was significantly larger for the human audio (mean 0.27, std 0.32) than for the animal audio (mean 0.12, std 0.31), as revealed by a two-tailed paired sample t -test $t(25) = 4.011, p < 0.001$. The three-way interaction indicates the differential influence of the sources varies across visual emotion. We therefore computed the difference between the congruent and incongruent conditions for every auditory source and visual emotion. Paired t -tests showed for both happy and fearful

body language a significant difference between the human congruent minus incongruent measure and the animal congruent minus incongruent measure.

Since a delayed reaction time task was used, no reaction time data were analyzed.

Discussion

The first aim of the present study was to investigate whether auditory information influences recognition of the emotion expressed in a simultaneously presented dynamic body expression. To test whether such crossmodal influence obtains, we presented video clips paired with nonverbal vocalizations and presented these stimuli with the instruction to categorize the emotion expressed by the body while ignoring the information provided by the auditory channel. Our results clearly indicate that recognition of body expressions is influenced by nonverbal vocal expressions. These findings are consistent with previous reports of crossmodal bias effects of vocal expressions on recognition of facial expressions, so far all using verbal stimuli (de Gelder et al., 1999; de Gelder & Vroomen, 2000; Ethofer et al., 2006; Massaro & Egan, 1996).

Our second aim was to investigate whether crossmodal influence is dependent on the perceived source of the auditory information or also obtains when different sources (human or animal sounds) have a similar signal function. Indeed, we find a clear influence of task irrelevant human voices on recognition of body language. However, the results also demonstrate that recognition of body language is influenced by environmental sounds. Happy body language is recognized better in combination with joyful bird songs, and recognized worse in combination with aggressive dog barks, compared to when the same happy body language is presented without auditory information. Human bodies are more intimately linked to human vocal expressions than animal vocalizations, which suggest that crossmodal influences are more probable in body-voice pairs, even if both can be perceived as carrying the same meaning, a typical example being danger signaling. The significant auditory emotion x source two-way interaction indicates that the impact of human vocalizations on the recognition of body language is larger than the impact of animal vocalizations. In view of the results of the pilot study which showed that human and animal vocalizations are recognized equally well, one may take this result as indicating that in general, human sounds influence

recognition of human body language to a greater extent than animal sounds. Such an interpretation would be consistent with views in the literature on the importance of semantic and cognitive factors in multisensory pairing. A more finely tuned comparison of the impact of both sources would need a more detailed balancing of both sources, for example on the basis of the variability in pitch and volume. Much as such controls are needed in future research, we would like to point out that controlling the physical dimensions does not settle questions on the role of semantic and cognitive factors affecting crossmodal bias (de Gelder & Bertelson, 2003). The nature of the privileged link between a facial or a bodily expression and auditory affective information produced by a person is at present not well understood. Similarly, comparisons between human sounds and the ones present in the environment have so far not been undertaken frequently. One recent suggestion is that the link between human face-body expressions and human vocalizations is based on premotor and motor structures in charge of producing the seen as well as the heard actions (Kohler *et al.*, 2002). This would indeed explain the special status of human vocalizations observed here. But clear evidence in support of this view is currently not available. On the other hand, if at present there were a body of evidence, as for example could be provided by brain imaging studies, in support of the notion that heard and seen emotional expressions activate similar brain areas, alternative explanations come to mind. In fact, seen and heard emotion expressions may both activate the same “affect program” as argued for example by Tomkins (1962, 1963) and later Ekman (1982). Known convergence of auditory and visual afferents on the amygdala support this view (Dolan, Morris, & de Gelder, 2001). The latter alternative can accommodate easily the similarity in emotional signal function between human and animal sounds without appeal to a perception/production link. The present study raises these questions as topics for future research in the relatively novel field which will need to address the issues raised for three decades concerning the links between seen and heard speech perception. In the same vein future research will address the question whether the crossmodal bias also obtains between a visual image and a written word instead of its sound referent. This is again a matter that has been investigated in the area of audiovisual speech and been answered negatively (Vroomen & de Gelder, 2000)

The crossmodal influence we observe here is slightly different depending on whether the bodily expressions are fearful or happy. A comparison of the AV-conditions with the V-condition yielded a performance increase in AV-congruent condition and a performance decline in AV-incongruent condition for happy bodily expressions. For the fearful body language, we observe only a performance decline in AV-incongruent condition. So we find

for the happy body language both a congruency and incongruency effect, but for the fearful body language, we find only an incongruence effect. The lack of a congruency effect for fearful body language cannot be explained by a ceiling effect given the results of the pilot data. We have currently no solid explanation for this differential crossmodal influence on the happy and fearful body language. An interesting topic for future research concerns the question whether the magnitude of emotional crossmodal influence differs between different emotions.

The results from the present study clearly indicate that crossmodal influences also occur even if both modalities are unmistakably produced by a different source. A relevant question would be what the conditions are for bimodal stimuli to be susceptible to crossmodal influences. Next to the obvious conditions of temporal and spatial congruence, animacy could play a role in the case of social stimuli. A recent event related potential (ERP) study compared brain waveforms when perceiving human faces paired with either a human burp, or a monkey scream or a squeaking door. Results pointed to animacy specific neural responses, next to species-specific brain waveforms (Puce, Epling, Thompson, & Carrick, 2007).

An important issue concerns the nature of the crossmodal influence. On the basis of a behavioral study, no direct inference can be made that the observed crossmodal influence has a perceptual basis. However, the instructions explicitly stated to base the emotional categorization solely on one modality (i.e. the visual), which is standard procedure in research dealing with conflicting multimodal inputs (Bertelson, 1998) and suggests an integrative perceptual process (de Gelder & Vroomen, 2000). Crossmodal integration of face-voice pairs seems unaffected by attentional resources (Vroomen, Driver, & de Gelder, 2001) and a study with event-related potentials (ERP) indicates a very early integration of emotional faces and voices (around 110 ms after stimulus onset) (Pourtois, de Gelder, Vroomen, Rossion, & Crommelinck, 2000). To examine the possible perceptual basis of a crossmodal bias effect with a behavioral paradigm, the ratings of unimodal stimuli in a pre-test could be compared with the ratings of a post-test, with repeated presentations of bimodal pairs between the pre-test and post-test. The presence of after-effects of the bimodal presentations on the post-test unimodal ratings would point to a perceptual influence of the auditory information. The present study indicates the occurrence of crossmodal influences of both human and animal vocalizations on the recognition of dynamic body language, but does not allow conclusions concerning the nature of the effects.

Ecological validity is an important factor in multisensory integration (de Gelder & Bertelson, 2003). Multimodal inputs reduce stimulus ambiguity and the brain has primarily evolved to maximize adaptiveness in the real world, and this is one of the reasons why we choose visual stimuli with high ecological validity, namely the performance of an everyday action in the context of a realistic situation.

Recent functional magnetic resonance imaging (fMRI) studies looked at the neural correlates of integrating emotional faces and voices (Dolan et al., 2001; Ethofer et al., 2006; Kreifelts, Ethofer, Grodd, Erb, & Wildgruber, 2007) and found increased activity in the left amygdala, when a fearful face was presented with a fearful voice (Dolan et al., 2001; Ethofer et al., 2006). The amygdala receive inputs from visual and auditory association cortex in the mammalian brain (McDonald, 1998) and its role in processing emotional stimuli is well established (see Zald, 2003 for a review). The amygdala therefore seems a primary candidate brain structure for integrating emotional information from different modalities.

References

- Atkinson, A. P., Dittrich, W. H., Gemmell, A. J., & Young, A. W. (2004). Emotion perception from dynamic and static body expressions in point-light and full-light displays. *Perception*, 33(6), 717-746.
- Bertelson, P. (1998). Starting from the ventriloquist: The perception of multimodal events. In M. Sabourin, I. Craik & M. Roberts (Eds.), *Advances in Psychological Science* (Vol. 2: Biological and Cognitive Aspects, pp. 419-439). Hove, UK: Lawrence Erlbaum.
- de Gelder, B. (2006). Towards the neurobiology of emotional body language. *Nature Reviews Neuroscience*, 7(3), 242-249.
- de Gelder, B., & Bertelson, P. (2003). Multisensory integration, perception and ecological validity. *Trends Cogn Sci*, 7(10), 460-467.
- de Gelder, B., Bocker, K. B., Tuomainen, J., Hensen, M., & Vroomen, J. (1999). The combined perception of emotion from voice and face: early interaction revealed by human electric brain responses. *Neurosci Lett*, 260(2), 133-136.
- de Gelder, B., & Vroomen, J. (2000). The perception of emotions by ear and by eye. *Cognition and Emotion*, 14(3), 289-311.

- Dolan, R. J., Morris, J. S., & de Gelder, B. (2001). Crossmodal binding of fear in voice and face. *Proc Natl Acad Sci U S A*, 98(17), 10006-10010.
- Ekman, P. (1982). *Emotion in the human face*. Cambridge: Cambridge University Press.
- Ethofer, T., Anders, S., Erb, M., Droll, C., Royen, L., Saur, R., et al. (2006). Impact of voice on emotional judgment of faces: an event-related fMRI study. *Hum Brain Mapp*, 27(9), 707-714.
- Grezes, J., Pichon, S., & de Gelder, B. (2007). Perceiving fear in dynamic body expressions. *Neuroimage*, 35(2), 959-967.
- Kohler, E., Keysers, C., Umiltà, M. A., Fogassi, L., Gallese, V., & Rizzolatti, G. (2002). Hearing sounds, understanding actions: action representation in mirror neurons. *Science*, 297(5582), 846-848.
- Kreifelts, B., Ethofer, T., Grodd, W., Erb, M., & Wildgruber, D. (2007). Audiovisual integration of emotional signals in voice and face: an event-related fMRI study. *Neuroimage*, 37(4), 1445-1456.
- Massaro, D. W., & Egan, P. B. (1996). Perceiving Affect from the Voice and the Face. *Psychonomic Bulletin and Review*, 3, 215-221.
- McDonald, A. J. (1998). Cortical pathways to the mammalian amygdala. *Prog Neurobiol*, 55(3), 257-332.
- Müller, J. P. (1840). *Handbuch der Physiologie des Menschen*. Coblenz: H. Ischer.
- Peelen, M. V., & Downing, P. E. (2007). The neural basis of visual body perception. *Nat Rev Neurosci*, 8(8), 636-648.
- Pourtois, G., de Gelder, B., Vroomen, J., Rossion, B., & Crommelinck, M. (2000). The time-course of intermodal binding between seeing and hearing affective information. *Neuroreport*, 11(6), 1329-1333.
- Puce, A., Epling, J. A., Thompson, J. C., & Carrick, O. K. (2007). Neural responses elicited to face motion and vocalization pairings. *Neuropsychologia*, 45(1), 93-106.

- Spreckelmeyer, K. N., Kutas, M., Urbach, T. P., Altenmuller, E., & Munte, T. F. (2006). Combined perception of emotion in pictures and musical sounds. *Brain Res*, 1070(1), 160-170.
- Tomkins, S. S. (1962). *Affect, imagery and consciousness: The positive affects*. New York: Springer Verlag.
- Tomkins, S. S. (1963). Affect, imagery consciousness: Vol. 2. The negative affects. In. New York: Springer verlag.
- Van den Stock, J., Righart, R., & de Gelder, B. (2007). Body expressions influence recognition of emotions in the face and voice. *Emotion*, 7(3), 487-494.
- Vroomen, J., & de Gelder, B. (2000). Crossmodal integration: a good fit is no criterion. *Trends Cogn Sci*, 4(2), 37-38.
- Vroomen, J., Driver, J., & de Gelder, B. (2001). Is cross-modal integration of emotional expressions independent of attentional resources? *Cogn Affect Behav Neurosci*, 1(4), 382-387.
- Welch, R. B., & Warren, D. H. (1986). Intersensory interactions. In K. R. Boff, L. Kaufman & J. P. Thomas (Eds.), *Handbook of perception and performance: Vol. 1. Sensory processes and perception* (pp. 25.21-25.36). New York: John Wiley and Sons.
- Zald, D. H. (2003). The human amygdala and the emotional evaluation of sensory stimuli. *Brain Res Brain Res Rev*, 41(1), 88-123.

Chapter 6: Instrumental music influences recognition of emotional body language⁶

Introduction

The movie ‘2001: A Space Odyssey’ is a landmark in the science-fiction genre. A classic scene shows a man-ape smashing a skeleton with a bone, while Richard Strauss’s *Also Sprach Zarathustra* blasts in the background. It is the combination of these visual and auditory inputs that results in a unique experience in the viewer.

Research on multisensory perception has a long history (Müller, 1840) and focussed on audiovisual speech (McGurk & MacDonald, 1976). However, multisensory research on emotional events is scarce and was until recently limited to investigations into the perception of facial and vocal expressions (e.g. de Gelder & Vroomen, 2000). In the latter type of studies two modalities are typically combined to create emotionally congruent and incongruent face-voice pairs and to provide a window into the integration process (de Gelder & Bertelson, 2003). Participants are instructed to rate the emotion in one of the two modalities while ignoring the other. The results have shown that recognition of the emotion in the target modality is typically influenced towards the emotion expressed in the task irrelevant modality.

In two recent studies we have taken this issue beyond facial expressions and investigated affective crossmodal influences in whole body expressions (Van den Stock, Grezes, & de Gelder, 2008). Here we investigated naturalistic actions that are part of everyday life and focussed on instrumental actions, like grasping and drinking. Our data showed that affective crossmodal effects occur with body-voice pairs, but also when body expressions are presented with animal vocalizations (Van den Stock et al., 2008). In the present study, we take the issue of affective crossmodal influence a step further and focus on bimodal stimuli that are not normally associated with each other, namely instrumental classical music and a person involved in an everyday action in a natural location.

⁶ This chapter is based on: Van den Stock, J., Peretz, I., Grèzes, J., & de Gelder, B. (2009). Instrumental music influences recognition of emotional body language. *Brain Topography*, 21(3-4), 216-220.

Method

Participants. Fifteen adult participants (7 male) were recruited through the local newspaper and were paid 20€. Mean age (SD) was 44.0 (12.6) years. Guidelines of the Declaration of Helsinki were followed.

Materials. Visual materials consisted of video recordings of 12 actors (6 male) (1998) who performed an everyday action (picking up a glass, drinking from it and putting it back on the table) and were shown in full body view. See Fig 1 for examples. They performed this action with different emotional expressions (anger, disgust, fear, happiness, sadness and neutral). Before the performance, they were briefed with a specific scenario. For example, the happy scenario specified that the glass contained the favourite drink of the actor. The scenarios for all emotions are shown in Table 1. 3000ms fragments were taken from the recordings and the faces of the actors were blurred. We used video editing software (Adobe Aftereffects 8.0) to track the trajectory of the face in the movie and we replaced the face by a blurred mask. In a pilot study, all edited stimuli were presented 4 times in random order to 14 participants. They were instructed to categorize the emotion expressed by the actor in a 6 alternative forced choice task (anger, disgust, fear, happiness, sadness and neutral). On the basis of these, we selected 10 happy videos (5 male) and 10 sad videos (5 male) that were all correctly recognized above 75%.

Table 1: scenarios provided to the actors to perform the action with different emotions

Emotion	Scenario
Anger	You just had an intense quarrel with a friend and you are very angry when you drink from the glass
Disgust	The glass contains a stomach-turning liquid and you are very disgusted
Fear	The content of the glass is extremely hot and you are afraid to drink it
Happiness	The glass contains your favourite drink and this makes you very happy
Sadness	You are returning from the funeral of a loved one and feel very sad, while you drink the glass
Neutral	Drink from the glass without any specific emotional state of mind

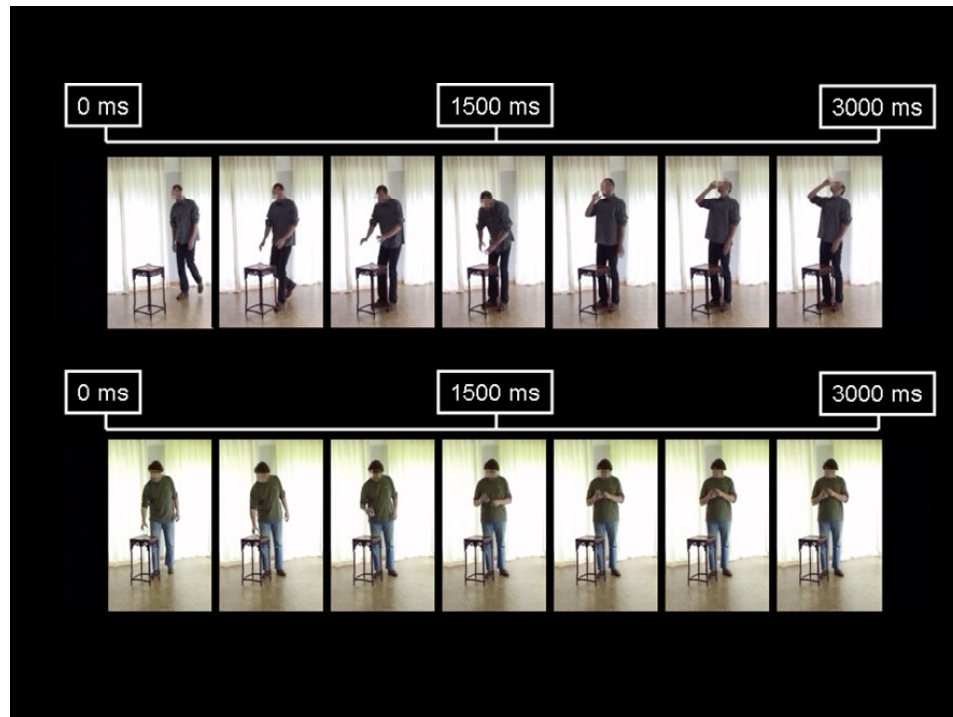


Figure 1. Examples of frames from the video clips. The top row shows frames from a happy video, the bottom row shows frames from a sad video

Auditory materials consisted of fragments from the classical repertoire that expressed a happy or a sad tone and are described and validated in Peretz et al. (1998). We selected a 3000ms fragment from 10 happy and 10 sad excerpts. Results in Peretz et al. (1998) show that distinct happy and sad ratings are already elicited after 500ms.

Procedure

Each of the 20 auditory stimuli was randomly paired once with a happy video and once with a sad video. This resulted in 40 unique bimodal stimuli of which 20 were congruent (e.g. happy audio paired with happy video) and 20 were incongruent (e.g. happy audio paired with sad video).

The experiment consisted of a visual (V), an auditory (A) and an audiovisual (AV) block. The order of blocks was randomized. A trial always started with presentation of a white fixation cross against a dark background shown for a variable duration of 1000-3000ms to reduce temporal predictability. This was followed by presentation of a stimulus (V, A or AV 3000ms) after which a dark screen with a white question mark was shown until a response was given. In the V-block, the stimuli consisted of the 20 videos that were randomly presented one by one while the participants were instructed to categorize the emotion expressed by the actor. In the A-block the 20 auditory stimuli were randomly presented under the instruction to categorize the emotion expressed by the music. In the AV-block, all 40 bimodal stimuli were randomly presented and the instructions explicitly stated to categorize as accurately and as fast as possible the emotion expressed by the body language of the actor. The on- and offset of the visual and auditory stimuli in the AV-block were synchronized. Auditory stimuli were delivered through pc speakers located on the left and right of the screen. The volume was set at a comfortable listening level.

Results

The results are displayed in Fig 2. We calculated the proportion happy responses for every condition in every block. A paired samples t-test showed there was no significant difference between the proportion happy responses in the happy audio condition and the proportion sad responses in the sad audio condition ($t(14) = .168, p < .869$), indicating both expressions are equally well recognized. A repeated measures ANOVA on the proportion happy responses with visual emotion (2 levels: happy and sad) and auditory emotion (3 levels: happy, sad and no audio) as within-subjects factors revealed a significant main effect of visual emotion ($F(1,14) = 506, p < .001$) and of auditory emotion ($F(2,28) = 4.734, p < .017$). The interaction was not significant ($F(2,28) = .278, p < .760$). The main effect of visual emotion indicates that the proportion happy responses is higher in the conditions with happy body language. This is to be expected since the task involved categorization of the emotion expressed by the body. Therefore, the main effect of visual emotion merely indicates that happy body expressions are more frequently rated as happy compared to sad body expressions. The main effect of auditory emotion indicates that the proportion happy responses differs according to auditory condition. Pairwise post-hoc comparisons (LSD corrected) on the main effect of auditory emotion showed that the proportion happy responses is significantly higher in the conditions with happy audio, compared to both sad audio

($p < .039$) and to no audio ($p < .005$). This means that both visual conditions, namely happy and sad dynamic whole body expressions are categorized more frequently as expressing happiness when they are presented simultaneously with happy music compared to when each of these visual conditions are presented with sad music or without auditory information.



Fig 2 Proportion 'happy' responses as a function of body expression and auditory information. Error bars represent 1 SEM.

The absence of a visual emotion \times auditory emotion interaction effect indicates that the effect of auditory information is of equal magnitude in both visual conditions. The results also show that the proportion happy responses is lowest in the condition with sad audio, but the difference with the other audio conditions was not significant ($p < .086$).

Since the task stated to respond when the question mark appeared, no reaction time data were analysed.

Discussion

The results show that emotional dynamic whole body expressions presented with happy music are recognized more as happy compared to when the same body expressions are paired with sad classical music or without auditory information. Even when instructions explicitly state to categorize the emotion expressed by the visual stimulus, the ratings are influenced towards the emotion expressed by the auditory stimulus. Our findings show that the influence of happy music is equally pronounced in both happy and sad body language. Moreover, body expressions presented with sad music are recognized more frequently as sad compared to when the bodies are presented in isolation or with happy music, but this effect is only marginally significant. The stronger influence of the auditory material on the happy body expressions compared to the sad expressions might be related to the level of intensity, valence and/or arousal expressed in the visual stimuli. Another possibility regards the matching of visual and auditory dynamics. It is not unlikely that congruence between changes in musical tempo and visual movement contributes to crossmodal influences. In previous studies, we have shown that whole body expressions of emotion can influence recognition of vocal emotional expressions (Van den Stock, Righart, & de Gelder, 2007), but also that whole body expressions are influenced by both human and animal vocalizations (Van den Stock et al., 2008). The rate of co-occurrence of body-voice pairs in natural circumstances is high, since both are produced by the same source. Presumably the perceptual system is well versed in the simultaneous processing of jointly produced or at least naturalistically co-occurring visual and auditory inputs and may therefore rely on specialized mechanisms for crossmodal binding (de Gelder & Bertelson, 2003). The combination of whole body expressions and animal vocalizations is less frequent and by this reasoning crossmodal influence between these stimulus categories is less evident. Still, simultaneous perception of a fearful body expression and a fear inducing dog bark can be perceived as one event, especially considered from an evolutionary perspective. However, the evolutionary significance of watching a person grasping an object while hearing instrumental classical music is less direct and less understood. The importance of the present findings lies in the fact that even multimodal inputs with no direct strong adaptive association can modulate the affective interpretation of clearly separate information streams. Nevertheless, instrumental music and body movements certainly occur frequently in dance, movies, social situations, etc. Our results show it is worth considering that the brain is

organized for maintaining these flexible associations. Music may also mimic prosodic cues that otherwise communicate emotion vocally or through ambient environmental sounds. Even if the effectiveness of music for conveying emotion is entirely a learned process shaped by culture, it is interesting that the brain has found a way to link music to emotion and furthermore to cross-modally link music and bodily cues.

The pilot validation study consisted of a 6 alternative forced choice design. The primary aim of this pilot study was to assess how well the stimuli expressed the target emotion and therefore we offered the participants a range of response alternatives. In the main experiment, we choose to administer a design with 2 response alternatives in keeping with the design of our previous experiments (Van den Stock et al., 2008; Van den Stock et al., 2007). The aim of the main experiment was to investigate crossmodal influence and we believe that a more limited number of response possibilities is preferable when making affective judgments in this context. Increasing the number of response alternatives may involve a higher appeal to more cognitive processes.

Despite the fact that the data from the pilot validation study show that the visual stimuli are easily recognizable when it comes to emotional categorization, one can not entirely exclude that the action in itself, i.e. drinking has an emotionally neutral association. For example it may be that drinking is associated with relief of thirst and is thereby biased towards a positive valence. However, the main interest of the present study concerns crossmodal influence elicited by the auditory information and this is measured by the difference between visual and audiovisual conditions. The primary focus of this study is the change between how congruent, incongruent and unimodal stimuli are categorized and the valence of the action itself (drinking) is equal in all the conditions.

One possible explanation for the observed effects might be that both visual and auditory emotional information elicit a similar affect program (Panksepp, 1998; Tomkins, 1962, 1963), which is neuro-anatomically supported by the involvement of premotor structures in perception of both body expressions (de Gelder, Snyder, Greve, Gerard, & Hadjikhani, 2004) and music (Minati et al., 2008).

An alternative but not incompatible explanation at the neuro-anatomical level implies a link between production and perception of emotional actions. Bimodal mirror neurons in monkey premotor and motor structures display an increased firing rate when an action is either

performed, seen or heard (Kohler et al., 2002). Indirect evidence from functional Magnetic Resonance Imaging (fMRI) supports the existence of a similar mirror (Grèzes, Armony, Rowe, & Passingham, 2003) and bimodal mirror (Lahav, Saltzman, & Schlaug, 2007) system in premotor cortex in humans. The latter study shows that the premotor cortex of non-musicians is more activated by listening to musical excerpts that they have recently learned to play on the piano than by music they have never played. Although this study does not focus on the affective features of the music, it indicates that more complex auditory stimuli like classical music activate right premotor structures in humans. We have shown previously that perception of emotional body language also activates right premotor structures (de Gelder et al., 2004; Grezes, Pichon, & de Gelder, 2007). These combined findings may provide a neuro-anatomical framework to explain the crossmodal effects observed in the present study.

However, the focus of the present study concerns the multimodal integration of *emotional information* and the role of the amygdala in processing emotional information has been well established (see Zald, 2003 for a review). Moreover, previous studies using face-voice pairs have shown that crossmodal binding of affective information involves the left amygdala (Dolan, Morris, & de Gelder, 2001; Ethofer et al., 2006) and this brain structure receives both visual and auditory inputs (McDonald, 1998). Therefore, this may be a critical brain region involved in the unique experience one has when watching the ape-man in '2001: A Space Odyssey'.

The results of our behavioural study do not allow to formulate hypotheses about the perceptual underpinnings of the observed effects. We used a similar paradigm as our previous study (Van den Stock et al., 2008) but investigation of the stage at which affective crossmodal influence occurs, i.e. either a visual-perceptual level, a semantic post-perceptual level or even a response selection level, requires the complimentary use of imaging techniques, preferably with a high temporal resolution like electroencephalography (EEG) or magnetoencephalography (MEG).

Another issue concerns what it is in a dynamic whole body expression that makes it happy or sad. Sadness is typically more associated with lower muscle tonus and slower movements, whereas happiness usually involves quick and rapid movements, mostly involving raising of the arms. The movement parameters that are related to emotional communication have been extensively described earlier (e.g. Argyle, 1988; Darwin, 1872).

Our study makes a beginning with exploring how music influences the message conveyed by body language. The different levels at which music and body language make contact and the neurofunctional basis of our embodied music experience are just some of the many questions to be addressed in future research.

References

- Argyle, M. (1988). *Bodily communication*. London: Methuen.
- Darwin, C. (1872). *The expression of the emotions in man and animals*. London: John Murray.
- de Gelder, B., & Bertelson, P. (2003). Multisensory integration, perception and ecological validity. *Trends Cogn Sci*, 7(10), 460-467.
- de Gelder, B., Snyder, J., Greve, D., Gerard, G., & Hadjikhani, N. (2004). Fear fosters flight: A mechanism for fear contagion when perceiving emotion expressed by a whole body. *Proc Natl Acad Sci U S A*, 101(47), 16701-16706.
- de Gelder, B., & Vroomen, J. (2000). The perception of emotions by ear and by eye. *Cognition and Emotion*, 14(3), 289-311.
- Dolan, R. J., Morris, J. S., & de Gelder, B. (2001). Crossmodal binding of fear in voice and face. *Proc Natl Acad Sci U S A*, 98(17), 10006-10010.
- Ethofer, T., Anders, S., Erb, M., Droll, C., Royen, L., Saur, R., et al. (2006). Impact of voice on emotional judgment of faces: an event-related fMRI study. *Hum Brain Mapp*, 27(9), 707-714.
- Grèzes, J., Armony, J. L., Rowe, J., & Passingham, R. E. (2003). Activations related to "mirror" and "canonical" neurones in the human brain: an fMRI study. *Neuroimage*, 18(4), 928-937.
- Grezes, J., Pichon, S., & de Gelder, B. (2007). Perceiving fear in dynamic body expressions. *Neuroimage*, 35(2), 959-967.
- Kohler, E., Keysers, C., Umiltà, M. A., Fogassi, L., Gallese, V., & Rizzolatti, G. (2002). Hearing sounds, understanding actions: action representation in mirror neurons. *Science*, 297(5582), 846-848.
- Lahav, A., Saltzman, E., & Schlaug, G. (2007). Action representation of sound: audiomotor recognition network while listening to newly acquired actions. *J Neurosci*, 27(2), 308-314.
- McDonald, A. J. (1998). Cortical pathways to the mammalian amygdala. *Prog Neurobiol*, 55(3), 257-332.
- McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature*, 264(5588), 746-748.

- Minati, L., Rosazza, C., D'Incerti, L., Pietrocini, E., Valentini, L., Scaioli, V., et al. (2008). fMRI/ERP of musical syntax: comparison of melodies and unstructured note sequences. *Neuroreport*, 19(14), 1381-1385.
- Müller, J. P. (1840). *Handbuch der Physiologie des Menschen*. Coblenz: H. Ischer.
- Panksepp, J. (1998). *Affective neuroscience: The foundation of human and animal emotions*. New York: Oxford University Press.
- Peretz, I., Gagnon, L., & Bouchard, B. (1998). Music and emotion: perceptual determinants, immediacy, and isolation after brain damage. *Cognition*, 68(2), 111-141.
- Tomkins, S. S. (1962). *Affect, imagery and consciousness: The positive affects*. New York: Springer Verlag.
- Tomkins, S. S. (1963). *Affect, imagery consciousness: Vol. 2. The negative affects*. New York: Springer verlag.
- Van den Stock, J., Grezes, J., & de Gelder, B. (2008). Human and animal sounds influence recognition of body language. *Brain Research*, 1242, 185-190.
- Van den Stock, J., Righart, R., & de Gelder, B. (2007). Body expressions influence recognition of emotions in the face and voice. *Emotion*, 7(3), 487-494.
- Zald, D. H. (2003). The human amygdala and the emotional evaluation of sensory stimuli. *Brain Res Brain Res Rev*, 41(1), 88-123.

Chapter 7: Moving and being moved. The relative importance of dynamical information for residual face processing in clinical populations and brain damaged patients⁷

Introduction

In encountering a person, what we are most easily conscious of is that the face gives us access to the person's identity. But at the same time the face provides many other kinds of information like gender, age, emotional expression, attractiveness, trust and the like. It is likely that some kinds of information are relatively better conveyed by moving than by static faces. Some of these typical face attributes, like for example identity or affect, are also conveyed by other stimuli than faces, for example whole bodies. And as is the case with faces, they may be conveyed by a still image as well as by its dynamic counterpart. Thus the relative importance of dynamic information is not an issue restricted to face recognition but is encountered just as well in investigations of object recognition in general.

Of course, it may very well be the case that not only faces are "special" but that also the perception of dynamic information is special in the case of faces. Very few focussed comparisons are available to answer this question because this is a challenging task. A proper comparison of face perception recognition abilities with other object perception and recognition abilities requires comparable task settings between the two object classes (Damasio, Damasio, & Van Hoesen, 1982; Damasio, Tranel, & Damasio, 1990; B. de Gelder, Bachoud-Levi, & Degos, 1998; Farah, 1990; Gauthier, Behrmann, & Tarr, 1999; Van den Stock, van de Riet, Righart, & de Gelder, 2008). Available comparisons have almost all used still images and this makes it all the more difficult to assess the relative importance of dynamic information for face perception.

⁷ This chapter is based on de Gelder, B., & Van den Stock, J. (in press). Moving and being moved. The relative importance of dynamical information for residual face processing in clinical populations and brain damaged patients. In Giese, M., Curio, C., & Bülthoff H.H. (Eds.), *Dynamic Faces: Insights from Experiments and Computation*. MIT Press

Face perception: some antecedents

The high salience of faces in everyday life is taken for granted and is reflected in the number of studies devoted to face recognition. Research targeting face recognition got a significant boost from the discovery of face specific deficits after brain damage reported by Bodamer (1947). Investigations into the functional properties of face processing began with the first neuropsychological studies of Yin (1969) reporting a strong inversion effect for faces and has been growing exponentially since the beginning of brain imaging studies of face recognition. More and more clinical cases are also reported this last decade with specific impairments in face recognition abilities. An overview of findings from functional Magnetic Resonance Imaging (fMRI) studies in these clinical cases can be found in Van den Stock et al. (2008) and an overview of EEG studies in Righart & de Gelder (2007).

The combined findings from behavioural, clinical and neuro-imaging studies are integrated in theoretical models of face perception, of which the model of Bruce & Young (1986) has been one of the most influential. Since then a few other models of face perception have been proposed (Haxby, Hoffman et al. 2000; Adolphs 2002; de Gelder, Frissen et al. 2003; Johnson 2005). They have moved our understanding ahead mainly by integrating new findings about face recognition deficits, the neurofunctional basis, category specificity, relative separation of subsystems like identity and expression, genetic basis, the importance of movement information and the contribution of real world and context elements. The central notion in contemporary models is that different aspects of face perception, like identity, expression and direction of gaze are processed in a brain network of which the different areas show relative functional specialization.

The neurofunctional basis of facial identity processing in neurologically intact individuals is reasonably well understood. Sergent & Signoret (1992) first described the middle lateral fusiform gyrus (FG) to be responsive to faces. Kanwisher and colleagues later dubbed this region the fusiform face area (FFA) (1997). The occipital face area (OFA) is another very important face sensitive area located in the inferior occipital gyrus (Gauthier, Skudlarski, Gore, & Anderson, 2000; Hoffman & Haxby, 2000; Puce, Allison, Asgari, Gore, &

McCarthy, 1996). While these areas have been related to identity processing the main area that comes into play when the face carries an emotional expression is the amygdala (AMG). The AMG plays a critical role in mediating *emotional* responses and *actions* (see Zald, 2003 for a review). Several studies support the notion that activity in FFA increases as a result from feedback from the AMG (e.g. Breiter, Etcoff et al. 1996; Vuilleumier, Richardson et al. 2004) and anatomical connections between amygdala and visual cortex have been established in primates (Amaral & Price, 1984; Freese & Amaral, 2006). Faces expressing emotions additionally modulate OFA activity (Rotshtein, Malach, Hadar, Graif, & Hendler, 2001). On the other hand, AMG driven threat-related modulations implicate also earlier visual areas such as V1 and other distant regions involved in social, cognitive, or somatic responses (e.g., STS, cingulate, or parietal areas) (Amaral 2002; Amaral, Behniea et al. 2003; Catani, Jones et al. 2003). The rapid activity and/or the involvement of posterior visual areas in normals have been related to coarse processing of salient stimuli in subcortical structures. Support for subcortical processing of salient stimuli of which facial expressions are a salient example is also provided by residual face perception in patients with striate cortex lesions (Morris, de Gelder et al. 2001; de Gelder 2005).

The brain basis of face perception in neurologically intact individuals: Perceiving movement

It needs no argument that in daily life, the faces we perceive and interact with are almost continuously in motion and our perceptual systems has therefore more experience with dynamic than with static faces. Movements generated by the complex musculature of the face or body make a substantial contribution to nonverbal communication. Moreover, there are several characteristics of a person that are almost exclusively revealed by its dynamic properties displayed in the face or body: looking at a photograph of Marlon Brando playing Don Corleone in the Godfather, results in an experience quite different from that of watching his whole performance in the scene in which he addresses the heads of the families. This difference is illustrative of the clear additive value that lies in the temporal unfolding of dynamic facial expressions.

Before developing this point though it is worth mentioning that using still images to probe face processing may have unique advantages to probe the neurofunctional basis in normal as well as in neurological patients. Static patterns get moving the mind as the brain processes the incoming still image by mapping it actively onto a representation that incorporates the movement and its temporal dynamics normally associated with this visual stimulus in the external world. Well known studies by e.g. Sheppard (1983) and by Freyd (1983) have shown convincingly that still images can be very fruitfully used to probe movement perception in the brain. Using still images of whole body expressions we observed activation in brain areas that are normally sensitive to movement like STS in human observers (B. de Gelder, Snyder, Greve, Gerard, & Hadjikhani, 2004) and macaques (B. de Gelder & Partan, 2009).

Although the importance of dynamic expressions and its interpretation by conspecifics has long been recognized in animal literature (Dawkins & Krebs, 1978), it is quite surprising there are still only very few neuro-imaging investigations with neurologically intact participants that used dynamic expressions. The dynamic information in facial expressions represents a specific kind of biological motion (Johansson, 1973). Therefore it is reasonable to expect that perceiving facial and bodily movements will activate areas known to be involved in movement perception like the hMT/V5 complex and in socially relevant movement perception like superior temporal sulcus (STS) (Bonda, Petrides, Ostry, & Evans, 1996; Grèzes & Decety, 2001; Grossman et al., 2000). Furthermore, socially relevant and emotionally laden movement is likely to involve AMG.

A few studies throw light on these issues but many open questions remain. For example, it is not known whether the neurofunctional basis of biological movement in faces and bodies is a special case of the more general ability for processing biological as contrasted with non biological movement. Alternatively, facial movement patterns that are specifically at the service of facial expressions may be a *sui generis* specialisation of the brain which only minimally overlaps with the neurofunctional mechanisms sustaining biological movement perception in general. The former possibility evokes the notion of a specialized speech module exclusively at the service of the analysis of visual speech. Liberman and colleagues developed the argument for such a phonetic gesture analysis module in the seventies and eighties. A review of the pro and contra arguments is provided in the volume dedicated to AI

Liberman. More recently, this approach to speech has been viewed as an example of action perception by researchers in the field of mirror neuron based action perception. But once relatively complex stimuli are considered, it remains unclear so far what the relation is between the movement and action perception (Pichon, de Gelder, & Grezes, 2009). Furthermore, the motor theory of speech perception was motivated by the ambition to start from but reach beyond the available linguistic description of phonetic features and define the set of motor primitives that may be at the basis of speech perception. Neither for the more general case of biological movement, nor for the specific ones of human facial movements do we have at present descriptive theories available. And studies of mirror neuron activation have so far been restricted to individual single actions that do not yet allow insight into action primitives. Possibly the analysis of facial motor patterns (FACS, Ekman) and bodily emotional motor patterns (BACS, de Gelder & van Boxtel) that implement emotional expressions may provide input for a future theory of emotional movement primitives.

With these caveats in mind, let us turn to available research. In a positron emission tomography (PET) study by Kilts et al. (2003) participants were presented angry, happy and neutral facial expressions and also non-face objects that were either static or dynamic. Perception of dynamic neutral faces compared to dynamic non-face objects triggered activity in AMG, STS and FG, but none of these areas were active when dynamic neutral faces were compared with static neutral faces. However, dynamic angry faces elicited more activity in these areas compared to static angry faces. This highlights the importance of emotional information conveyed by facial expressions in the comparison between dynamic and static faces. The increased recruitment of AMG, STS and FG in dynamic facial expressions might be specific for expressions with a negative valence, since there was no difference in these areas between dynamic and static displays of happy faces. Similar findings are reported with fMRI data: dynamic facial expressions (neutral, fear and anger) yielded more activity than static emotional faces in AMG and FG (LaBar, Crupain, Voyvodic, & McCarthy, 2003). An overview of currently available functional imaging studies with dynamic facial expressions is given in Table 1. The general findings show that comparisons between dynamic faces and dynamic non-face stimuli typically activate brain areas known already to be involved in perception of static faces. Taken at face value this result suggests that the difference in brain basis between seeing still and dynamic faces is quantitative rather than qualitative. However, a

more focussed comparison between dynamic and static faces shows a less clear picture, and the contrast becomes stronger when emotional expressions are part of the comparison.

Table 1. Overview of dynamic facial expression neuroimaging studies. D=dynamic; S=static; Fang=angry face; Fdis=disgusted face; Ffea=fearful face; Fneu=neutral face; Fhap=happy face; Femo=emotional face; Fall=all faces; nonF=non-face; R=rest; (s)=synthetic; (l) line drawing; Bang=angry body; Bfea=fearful body; Bneu=neutral body; Ball= all bodies. ¹The conditions reported involve the motoric expression of the stimuli, not the semantic content of the stories told by the actor. ²No modulation of AMG activity by emotional content of faces

	Method	Stimuli	Task	Contrast	AMG	FG	STS	OFA
Kilts (2003)	PET	FneuD	Emotion intensity rating	FneuD>DnonF	X	X	X	
		FangD		FneuD>FneuS				X
		FhapD		FangD>FangS	X	X	X	
		FneuS		FhapD>FhapS				X
		FfeaS						
		FangS						
		nonFD						
		nonFS						
LaBar (2003)	fMRI	FneuD	Category classification	FneuD>FneuS	X	X		
		FangD		FemoD>FemoS	X	X		
		FfeaD		FangD>FangS		X		
		FneuS		FfeaD>FfeaS	X	X		
		FfeaS						
		FfeaS						
Puce (2003)	fMRI	FneuD	Passive viewing	[FneuD+FneuD(l)]>nonFD		X	X	
		FneuD(l)						
		nonFD						
Decety (2003)	PET	FneuD ¹	Mood rating	FhapD>FneuD				
		FhapD ¹		FsadD>FneuD				
		FsadD ¹						
Wicker (2003)	fMRI	FneuD	Passive viewing	FdisD>FneuD		X	X	X
		FdisD		FhapD>FneuD		X		
		FhapD						
Sato (2004)	fMRI	FfeaD	Passive viewing	FfeaD>FfeaS	X	X	X	X
		FhapD		FfeaD>nonFD		X	X	X
		FfeaS		FhapD>FhapS		X	X	X

		FangS nonFD		FhapD>nonFD	X	X	X	X
Wheaton (2004)	fMRI	FneuD FneuS	Passive viewing	FneuD>FneuS		X	X	
Grosbras (2006)	fMRI	FneuD FangD nonFD	Passive viewing	FneuD>nonFD FangD>nonFD	X X	X X	X X	X X
van der Gaag (2007)a	fMRI	FneuD FdisD FfeaD FhapD	Passive viewing	FalID>R	X	X		
van der Gaag (2007)b	fMRI	FneuD FdisD FfeaD FhapD nonFD	Passive viewing Discrimination	FalID>nonFD FalID>nonFD	X ² X ²			
Thompson (2007)	fMRI	FneuD(s) nonFD	Speed change detection	FneuD>nonFD			X	
Kret (2009)	fMRI	FangD FfeaD FneuD BangD BfeaD BneuD	Oddball detection	FalID>BalID FemoD>FneuD	X	X	X	

In a recent study we investigated the neural correlates of perceiving dynamic face images using a design built on a close comparison of face videos with body videos. To arrive at a better view of dynamic neutral vs. emotional (fearful and angry) facial expressions we used both categories and compared each with its counterpart (Kret, Grezes, Pichon, & de Gelder, submitted). The *face versus body* comparison showed activation in the AMG and hippocampus. Dynamic emotional faces compared to dynamic neutral faces yielded more activity in FG and STS. We found no emotional modulation of the AMG by dynamic emotional compared to neutral faces, a result that is consistent with a study that focussed on amygdala activation (van der Gaag, Minderaa, & Keysers, 2007a).

Neurophysiological studies in monkeys

Single cell recordings in monkeys have shown that cells in IT and STS are responsive to different aspects of face perception (Baylis, Rolls, & Leonard, 1987; C. Bruce, Desimone, & Gross, 1981; De Souza, Eifuku, Tamura, Nishijo, & Ono, 2005; Desimone, Albright, Gross, & Bruce, 1984; D.I. Perrett, Hietanen, Oram, & Benson, 1992; D. I. Perrett, Rolls, & Caan, 1982), including emotional expression (Hasselmo, Rolls, & Baylis, 1989). However, the use of dynamic face-stimuli in neurophysiological monkey studies is rare. Evidence exists of neurons that are sensitive to specific movements of the head (Hasselmo, Rolls, Baylis, & Nalwa, 1989), and dynamic whole body expressions (Jellema, Maassen, & Perrett, 2004; Jellema & Perrett, 2003; Oram & Perrett, 1996). One neurophysiological study reported neurons in monkey STS that are sensitive to facial dynamics like closing the eyes (Hasselmo, Rolls, Baylis et al., 1989). Other cells have been described that are sensitive to threatening open mouths (D.I. Perrett & Mistlin, 1990).

Neurons that are selective for the appearance and motions of conspecifics have been described in the temporal lobes of macaque monkeys. Cells in the superior temporal sulcus (STS) of the macaque brain react to important social signals like threatening expressions. Other cells have been found to be selective for dynamic components of expressions such as raising the eyebrows, head and body movements (see Allison, Puce, & McCarthy, 2000 for a review).

Similarly in humans, the STS was found to be active following social information when dynamic images were used (see Table 1). Neurons in the amygdala have also been reported to be responsive to social information in monkeys (Brothers & Ring, 1993; Brothers, Ring, & Kling, 1990).

Visual object agnosia and face agnosia or prosopagnosia

Prosopagnosia was first reported by Bodamer (1947). The deficit involves recognition of personal identity but not of facial expression and this dissociation has long been the cornerstone of the models of face processing in the neuropsychological literature of the last two decades and is at the basis of the face recognition model of Bruce & Young (1986). The typical complaint of a prosopagnosic regards the inability to recognize a person by the face.

This symptom is far more pronounced than the phenomenon everybody sometimes experiences when have trouble remembering from where or how they know a certain face. Prosopagnosics can even have difficulties recognizing the persons they are very close to like their direct family members.

The neural correlates of face deficits

The focus on finding the neural correlate of the physically defined face category raised the expectation that patients with face recognition deficits would show lesions or anomalous activation in the normal face areas. This has not always turned out to be the case as shown by some recent patient studies using brain imaging (e.g. Steeves et al., 2006). When we turn to developmental prosopagnosia (DP), the situation is not clearer. Investigations into the neuro-functional correlates of DP with fMRI have yielded inconsistent results. Several studies reported increased activity for perceiving faces compared to non-face stimuli in the well known face areas FFA and OFA (Avidan, Hasson, Malach, & Behrmann, 2005; Degutis, Bentin, Robertson, & D'Esposito, 2007; Hasson, Avidan, Deouell, Bentin, & Malach, 2003; Williams, Berberovic, & Mattingley, 2007) whereas the first fMRI-study including a DP case by Hadjikhani and de Gelder (2002) and a more recent study (Bentin, Degutis, D'Esposito, & Robertson, 2007) found no face-specific activation in these areas. These findings suggest that intact functioning of the FFA and IOG are necessary, but not sufficient for successful face recognition. An important issue concerns the emotional information contained by the perceived faces. Recently, we observed reduced activation levels in the FFA of three developmental prosopagnosics compared to control subjects when looking at neutral faces. However, there was no difference between both groups in the activation level of the FFA when the faces they viewed expressed either a happy or a fearful emotion (Van den Stock et al., 2008). In the same study, we investigated the neural correlates of perceiving neutral and emotional whole body expressions and the results showed that in prosopagnosics, perception of bodies is associated with increased activation in face areas and perception of faces elicits activity in body areas. Whole body expressions are quite eligible as control stimulus condition for faces, since they are comparable to faces on a number of variables, for instance, ability to express emotional information, gender, age, familiarity.

Still vs. dynamic face images in patient studies

Impairments in recognizing emotion or identity in facial expressions have been reported in a variety of syndromes like Huntington's disease (Sprengelmeyer et al., 1996), Wilson's disease (Wang, Hoosain, Yang, Meng, & Wang, 2003), Urbach-Wiethe disease (Adolphs, Tranel, Damasio, & Damasio, 1994), Parkinson's disease (Sprengelmeyer et al., 2003), autism spectrum disorder (see Sasson, 2006 for a review), obsessive-compulsive disorder (Sprengelmeyer et al., 1997), schizophrenia (see Mandal, Pandey, & Prasad, 1998 for a review), Alzheimer's disease (Hargrave, Maddock, & Stone, 2002), semantic dementia (Bozeat, Lambon Ralph, Patterson, Garrard, & Hodges, 2000), attention deficit hyperactivity disorder (Singh et al., 1998), amyotrophic lateral sclerosis (Zimmerman, Eslinger, Simmons, & Barrett, 2007) and frontotemporal dementia (Lavenex, Pasquier, Lebert, Petit, & Van der Linden, 1999). However, the bulk of these studies is based on the use of static stimuli and recognition of static facial expressions urges a larger effort of the brain compared to dynamic expressions since the brain has to account for the missing information of temporal dynamics. It is therefore not surprising that several patient studies found superior recognition of dynamic facial expressions compared to static expressions (Tomlinson, Jones, Johnston, Meaden, & Wink, 2006).

One study with a prosopagnosic reported impaired identity recognition of static face pictures, but not of dynamic faces (Steede, Tree, & Hole, 2007), a pattern that was not compatible with a similar previous study (Lander, Humphreys, & Bruce, 2004).

As far as recognition of facial speech expressions is concerned, we tested a patient with prosopagnosia using still images of facial expressions as well as dynamic videos (Beatrice de Gelder & Vroomen, 1998). Her performance with still facial expressions was poor but improved significantly when short videos were shown instead. The same pattern was observed in another prosopagnosic patient using point light displays of emotional face expressions (Humphreys, Donnelly, & Riddoch, 1993).

Being moved by still images

It is often assumed that dynamic stimuli are easier to decode than still images and the most frequent argument is that dynamic images are more natural or more ecological and thereby more representative of the visual input the brain has evolved for. As we already pointed out, comparisons are complicated by the simple fact that dynamic stimuli contain much more information than do still images. On the other hand, there are arguments about the specificity of movement perception that speak against a simple comparison that takes higher information content of dynamic images into account. One of these is the fact that there are known cases in the neuropsychological literature of movement perception disorders. One of the best ones is Zihl's patient with bilateral lesions to V5. This patient had a severe movement perception deficit but had no difficulty with recognizing people by the face and was not prosopagnosic. She was also able to read speech from static face images, but could not perceive speech from dynamic images (Campbell, Zihl, Massaro, Munhall, & Cohen, 1997). The reverse pattern was observed in a patient with lesions in V4 (Humphreys et al., 1993).

A convergent argument to which we have already alluded several times above in favour of a nuance to the distinction between still and dynamic images is that studies using still images have reported activation in motor and premotor areas. This clearly means the brain does not need to be shown movement for perceiving it. As a matter of fact, using still images may provide a tool for assessing the brains' perceptual abilities beyond the strictly physically present information.

Face perception in hemianope patients

Of particular interest for understanding the neurofunctional basis of facial movement perception are patients with damage to primary visual cortex. Previous studies of such rare cases have illustrated the extent of residual movement vision that does not depend on intact V1. Interestingly, movement perception with vs. without awareness is correlated with different stimulus properties (for a review, see Weiskrantz, 2009).

In our first investigation of residual vision of hemianope patients we used both still images and short video clips of faces and we found that only the video clips triggered a reliable recognition of facial expressions in the blind field. This suggested that the presence of movement may be a necessary condition for affective blindsight (B. de Gelder, Vroomen, Pourtois, & Weiskrantz, 1999). However in subsequent studies we used EEG and later fMRI measurements and found clear evidence that still images were processed also (Rossion, de Gelder et al. 2000; Morris, de Gelder et al. 2001; de Gelder, Pourtois et al. 2002). Returning to a more sensitive behavioural paradigm than direct guessing by using the redundant target effect based on the advantage derived from summation across the two hemifields, here the sighted and the blind one, we showed later that still face images are still reliably recognized (de Gelder, Pourtois et al. 2002; de Gelder, Morris et al. 2005). In a recent study we report that the presence of still facial and bodily images triggers muscular movements that can be measured by EMG (electromyography). These facial movements reflect the specific emotion expressed in the unseen stimulus, independently of whether it is a face or a body and have shorter latencies when triggered by an unseen than by a seen stimulus. But at no time are the subjects aware of the unseen stimulus or of their motor reaction to it (Tamietto et al., submitted).

Conclusion

The human perceptual system is eminently tuned to information provided by movement in the environment. But the corollary of this that when dealing with still images the brain will automatically represent the dynamic information that is not strictly speaking present in the stimulus. Perceptual deficits, either congenital or as a consequence of brain damage in the normally developed brain, challenge our current understanding of the neurofunctional basis of movement perception. On the one hand there is little doubt that moving images provide more and richer information which other things being equal may make it easier to access information for brains and perceptual systems weakened by disease. On the other hand, there is also little doubt that in the developed brain a certain division of labour underpins fluent perceptual abilities. To approach this neuronal division of labour as exclusively a matter of specialized face, movement or emotion modules may hamper our understanding of these perceptual abilities and of the active role of the perceptual system in dealing with stimuli.

References

- Adolphs, R., Tranel, D., Damasio, H., & Damasio, A. (1994). Impaired recognition of emotion in facial expressions following bilateral damage to the human amygdala. *Nature*, 372(6507), 669-672.
- Allison, T., Puce, A., & McCarthy, G. (2000). Social perception from visual cues: role of the STS region. *Trends Cogn Sci*, 4(7), 267-278.
- Amaral, D. G., & Price, J. L. (1984). Amygdalo-cortical projections in the monkey (*Macaca fascicularis*). *J Comp Neurol*, 230(4), 465-496.
- Avidan, G., Hasson, U., Malach, R., & Behrmann, M. (2005). Detailed exploration of face-related processing in congenital prosopagnosia: 2. Functional neuroimaging findings. *J Cogn Neurosci*, 17(7), 1150-1167.
- Baylis, G. C., Rolls, E. T., & Leonard, C. M. (1987). Functional subdivisions of the temporal lobe neocortex. *J Neurosci*, 7(2), 330-342.
- Bentin, S., Degutis, J. M., D'Esposito, M., & Robertson, L. C. (2007). Too Many Trees to See the Forest: Performance, Event-related Potential, and Functional Magnetic Resonance Imaging Manifestations of Integrative Congenital Prosopagnosia. *J Cogn Neurosci*, 19(1), 132-146.
- Bodamer, J. (1947). Die prosop-Agnosie. *Archiv fur Psychiatrie und Nervenkrankheiten*, 179, 6-53.
- Bonda, E., Petrides, M., Ostry, D., & Evans, A. (1996). Specific involvement of human parietal systems and the amygdala in the perception of biological motion. *J Neurosci*, 16(11), 3737-3744.
- Bozeat, S., Lambon Ralph, M. A., Patterson, K., Garrard, P., & Hodges, J. R. (2000). Non-verbal semantic impairment in semantic dementia. *Neuropsychologia*, 38(9), 1207-1215.
- Brothers, L., & Ring, B. (1993). Mesial temporal neurons in the macaque monkey with responses selective for aspects of social stimuli. *Behav Brain Res*, 57(1), 53-61.

- Brothers, L., Ring, B., & Kling, A. (1990). Response of neurons in the macaque amygdala to complex social stimuli. *Behav Brain Res*, 41(3), 199-213.
- Bruce, C., Desimone, R., & Gross, C. G. (1981). Visual properties of neurons in a polysensory area in superior temporal sulcus of the macaque. *J Neurophysiol*, 46(2), 369-384.
- Bruce, V., & Young, A. (1986). Understanding face recognition. *Br J Psychol*, 77 (Pt 3), 305-327.
- Campbell, R., Zihl, J., Massaro, D., Munhall, K., & Cohen, M. M. (1997). Speechreading in the akinetopsic patient, L.M. *Brain*, 120 (Pt 10), 1793-1803.
- Damasio, A. R., Damasio, H., & Van Hoesen, G. W. (1982). Prosopagnosia: anatomic basis and behavioral mechanisms. *Neurology*, 32(4), 331-341.
- Damasio, A. R., Tranel, D., & Damasio, H. (1990). Face agnosia and the neural substrates of memory. *Annu Rev Neurosci*, 13, 89-109.
- Dawkins, R., & Krebs, J. R. (1978). Animals' signals: information or manipulation. In J. R. Krebs & N. B. Davies (Eds.), *Behavioural Ecology: An Evolutionary Approach*. (pp. 282-309). Oxford: Blackwell Scientific Publications.
- de Gelder, B., Bachoud-Levi, A. C., & Degos, J. D. (1998). Inversion superiority in visual agnosia may be common to a variety of orientation polarised objects besides faces. *Vision Res*, 38(18), 2855-2861.
- de Gelder, B., & Partan, S. (2009). The neural basis of perceiving emotional bodily expressions in monkeys. *Neuroreport*.
- de Gelder, B., Snyder, J., Greve, D., Gerard, G., & Hadjikhani, N. (2004). Fear fosters flight: A mechanism for fear contagion when perceiving emotion expressed by a whole body. *Proc Natl Acad Sci U S A*, 101(47), 16701-16706.
- de Gelder, B., & Vroomen, J. (1998). Impairment of speech-reading in prosopagnosia. *Speech Communication*, 26(1-2), 89-96.
- de Gelder, B., Vroomen, J., Pourtois, G., & Weiskrantz, L. (1999). Non-conscious recognition of affect in the absence of striate cortex. *Neuroreport*, 10(18), 3759-3763.

- De Souza, W. C., Eifuku, S., Tamura, R., Nishijo, H., & Ono, T. (2005). Differential characteristics of face neuron responses within the anterior superior temporal sulcus of macaques. *J Neurophysiol*, 94(2), 1252-1266.
- Decety, J., & Chaminade, T. (2003). Neural correlates of feeling sympathy. *Neuropsychologia*, 41(2), 127-138.
- Degutis, J. M., Bentin, S., Robertson, L. C., & D'Esposito, M. (2007). Functional Plasticity in Ventral Temporal Cortex following Cognitive Rehabilitation of a Congenital Prosopagnosic. *J Cogn Neurosci*, 19(11), 1790-1802.
- Desimone, R., Albright, T. D., Gross, C. G., & Bruce, C. (1984). Stimulus-selective properties of inferior temporal neurons in the macaque. *J Neurosci*, 4(8), 2051-2062.
- Farah, M. (1990). *Visual agnosia: Disorders of visual recognition and what they tell us about normal vision*. Cambridge: MIT Press.
- Freese, J. L., & Amaral, D. G. (2006). Synaptic organization of projections from the amygdala to visual cortical areas TE and V1 in the macaque monkey. *J Comp Neurol*, 496(5), 655-667.
- Freyd, J. J. (1983). The mental representation of movement when static stimuli are viewed. *Percept Psychophys*, 33(6), 575-581.
- Gauthier, I., Behrmann, M., & Tarr, M. J. (1999). Can face recognition really be dissociated from object recognition? *J Cogn Neurosci*, 11, 349-370.
- Gauthier, I., Skudlarski, P., Gore, J. C., & Anderson, A. W. (2000). Expertise for cars and birds recruits brain areas involved in face recognition. *Nature Neuroscience*, 3(2), 191-197.
- Grèzes, J., & Decety, J. (2001). Functional anatomy of execution, mental simulation, observation, and verb generation of actions: a meta-analysis. *Hum Brain Mapp*, 12(1), 1-19.
- Grosbras, M. H., & Paus, T. (2006). Brain networks involved in viewing angry hands or faces. *Cereb Cortex*, 16(8), 1087-1096.

- Grossman, E., Donnelly, M., Price, R., Pickens, D., Morgan, V., Neighbor, G., et al. (2000). Brain areas involved in perception of biological motion. *J Cogn Neurosci*, 12(5), 711-720.
- Hadjikhani, N., & de Gelder, B. (2002). Neural basis of prosopagnosia: an fMRI study. *Hum Brain Mapp*, 16(3), 176-182.
- Hargrave, R., Maddock, R. J., & Stone, V. (2002). Impaired recognition of facial expressions of emotion in Alzheimer's disease. *J Neuropsychiatry Clin Neurosci*, 14(1), 64-71.
- Hasselmo, M. E., Rolls, E. T., & Baylis, G. C. (1989). The role of expression and identity in the face-selective responses of neurons in the temporal visual cortex of the monkey. *Behav Brain Res*, 32(3), 203-218.
- Hasselmo, M. E., Rolls, E. T., Baylis, G. C., & Nalwa, V. (1989). Object-centered encoding by face-selective neurons in the cortex in the superior temporal sulcus of the monkey. *Exp Brain Res*, 75(2), 417-429.
- Hasson, U., Avidan, G., Deouell, L. Y., Bentin, S., & Malach, R. (2003). Face-selective activation in a congenital prosopagnosic subject. *Journal of Cognitive Neuroscience*, 15(3), 419-431.
- Hoffman, E. A., & Haxby, J. V. (2000). Distinct representations of eye gaze and identity in the distributed human neural system for face perception. *Nat Neurosci*, 3(1), 80-84.
- Humphreys, G. W., Donnelly, N., & Riddoch, M. J. (1993). Expression is computed separately from facial identity, and it is computed separately for moving and static faces: neuropsychological evidence. *Neuropsychologia*, 31(2), 173-181.
- Jellema, T., Maassen, G., & Perrett, D. I. (2004). Single cell integration of animate form, motion and location in the superior temporal cortex of the macaque monkey. *Cereb Cortex*, 14(7), 781-790.
- Jellema, T., & Perrett, D. I. (2003). Cells in monkey STS responsive to articulated body motions and consequent static posture: a case of implied motion? *Neuropsychologia*, 41(13), 1728-1737.
- Johansson, G. (1973). Visual perception of biological motion and a model for its analysis. *Perception & Psychophysics*, 14, 201-211.

- Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: a module in human extrastriate cortex specialized for face perception. *J Neurosci*, 17(11), 4302-4311.
- Kilts, C. D., Egan, G., Gideon, D. A., Ely, T. D., & Hoffman, J. M. (2003). Dissociable neural pathways are involved in the recognition of emotion in static and dynamic facial expressions. *Neuroimage*, 18(1), 156-168.
- Kret, M. E., Pichon, S., Grèzes, J., & de Gelder, B. (submitted). Gender specific brain activations in perceiving threat from dynamic faces and bodies.
- LaBar, K. S., Crupain, M. J., Voyvodic, J. T., & McCarthy, G. (2003). Dynamic perception of facial affect and identity in the human brain. *Cereb Cortex*, 13(10), 1023-1033.
- Lander, K., Humphreys, G., & Bruce, V. (2004). Exploring the role of motion in prosopagnosia: recognizing, learning and matching faces. *Neurocase*, 10(6), 462-470.
- Lavenex, I., Pasquier, F., Lebert, F., Petit, H., & Van der Linden, M. (1999). Perception of emotion in frontotemporal dementia and Alzheimer disease. *Alzheimer Dis Assoc Disord*, 13(2), 96-101.
- Mandal, M. K., Pandey, R., & Prasad, A. B. (1998). Facial expressions of emotions and schizophrenia: a review. *Schizophr Bull*, 24(3), 399-412.
- Oram, M. W., & Perrett, D. I. (1996). Integration of form and motion in the anterior superior temporal polysensory area (STPa) of the macaque monkey. *J Neurophysiol*, 76(1), 109-129.
- Perrett, D. I., Hietanen, J. K., Oram, M. W., & Benson, P. J. (1992). Organization and functions of cells responsive to faces in the temporal cortex. *Philos Trans R Soc Lond B Biol Sci*, 335(1273), 23-30.
- Perrett, D. I., & Mistlin, A. J. (1990). Perception of facial attributes. In W. C. Stebbins & M. A. Berkley (Eds.), *Comparative perception, complex signals* (Vol. II, pp. 187-215). New York: John Wiley.
- Perrett, D. I., Rolls, E. T., & Caan, W. (1982). Visual neurones responsive to faces in the monkey temporal cortex. *Exp Brain Res*, 47(3), 329-342.

- Pichon, S., de Gelder, B., & Grezes, J. (2009). Two different faces of threat. Comparing the neural systems for recognizing fear and anger in dynamic body expressions. *Neuroimage*, 47(4), 1873-1883.
- Puce, A., Allison, T., Asgari, M., Gore, J. C., & McCarthy, G. (1996). Differential sensitivity of human visual cortex to faces, letterstrings, and textures: a functional magnetic resonance imaging study. *J Neurosci*, 16(16), 5205-5215.
- Puce, A., Syngienotis, A., Thompson, J. C., Abbott, D. F., Wheaton, K. J., & Castiello, U. (2003). The human temporal lobe integrates facial form and motion: evidence from fMRI and ERP studies. *Neuroimage*, 19(3), 861-869.
- Righart, R., & de Gelder, B. (2007). Impaired face and body perception in developmental prosopagnosia. *Proc Natl Acad Sci U S A*.
- Rotshtein, P., Malach, R., Hadar, U., Graif, M., & Hendler, T. (2001). Feeling or features: different sensitivity to emotion in high-order visual cortex and amygdala. *Neuron*, 32(4), 747-757.
- Sasson, N. J. (2006). The development of face processing in autism. *J Autism Dev Disord*, 36(3), 381-394.
- Sato, W., Kochiyama, T., Yoshikawa, S., Naito, E., & Matsumura, M. (2004). Enhanced neural activity in response to dynamic facial expressions of emotion: an fMRI study. *Brain Res Cogn Brain Res*, 20(1), 81-91.
- Sergent, J., & Signoret, J. L. (1992). Functional and anatomical decomposition of face processing: evidence from prosopagnosia and PET study of normal subjects. *Philos Trans R Soc Lond B Biol Sci*, 335(1273), 55-61; discussion 61-52.
- Shepard, R. N., & Zare, S. L. (1983). Path-guided apparent motion. *Science*, 220(4597), 632-634.
- Singh, S. D., Ellis, C. R., Winton, A. S., Singh, N. N., Leung, J. P., & Oswald, D. P. (1998). Recognition of facial expressions of emotion by children with attention-deficit hyperactivity disorder. *Behav Modif*, 22(2), 128-142.

- Sprengelmeyer, R., Young, A. W., Calder, A. J., Karnat, A., Lange, H., Homberg, V., et al. (1996). Loss of disgust. Perception of faces and emotions in Huntington's disease. *Brain*, 119 (Pt 5), 1647-1665.
- Sprengelmeyer, R., Young, A. W., Mahn, K., Schroeder, U., Woitalla, D., Buttner, T., et al. (2003). Facial expression recognition in people with medicated and unmedicated Parkinson's disease. *Neuropsychologia*, 41(8), 1047-1057.
- Sprengelmeyer, R., Young, A. W., Pundt, I., Sprengelmeyer, A., Calder, A. J., Berrios, G., et al. (1997). Disgust implicated in obsessive-compulsive disorder. *Proc R Soc Lond B Biol Sci*, 264(1389), 1767-1773.
- Steede, L. L., Tree, J. J., & Hole, G. J. (2007). I can't recognize your face but I can recognize its movement. *Cogn Neuropsychol*, 24(4), 451-466.
- Steeves, J. K., Culham, J. C., Duchaine, B. C., Pratesi, C. C., Valyear, K. F., Schindler, I., et al. (2006). The fusiform face area is not sufficient for face recognition: evidence from a patient with dense prosopagnosia and no occipital face area. *Neuropsychologia*, 44(4), 594-609.
- Thompson, J. C., Hardee, J. E., Panayiotou, A., Crewther, D., & Puce, A. (2007). Common and distinct brain activation to viewing dynamic sequences of face and hand movements. *Neuroimage*, 37(3), 966-973.
- Tomlinson, E. K., Jones, C. A., Johnston, R. A., Meaden, A., & Wink, B. (2006). Facial emotion recognition from moving and static point-light images in schizophrenia. *Schizophr Res*, 85(1-3), 96-105.
- Van den Stock, J., van de Riet, W. A., Righart, R., & de Gelder, B. (2008). Neural correlates of perceiving emotional faces and bodies in developmental prosopagnosia: an event-related fMRI-study. *PLoS ONE*, 3(9), e3195.
- van der Gaag, C., Minderaa, R. B., & Keysers, C. (2007a). The BOLD signal in the amygdala does not differentiate between dynamic facial expressions. *Soc Cogn Affect Neurosci*, 2(2), 93-103.
- van der Gaag, C., Minderaa, R. B., & Keysers, C. (2007b). Facial expressions: what the mirror neuron system can and cannot tell us. *Soc Neurosci*, 2(3-4), 179-222.

- Wang, K., Hoosain, R., Yang, R. M., Meng, Y., & Wang, C. Q. (2003). Impairment of recognition of disgust in Chinese with Huntington's or Wilson's disease. *Neuropsychologia*, 41(5), 527-537.
- Weiskrantz, L. (2009). *Blindsight: a case study spanning 35 years and new developments*. Oxford: Oxford University Press.
- Wheaton, K. J., Thompson, J. C., Syngieniotis, A., Abbott, D. F., & Puce, A. (2004). Viewing the motion of human body parts activates different regions of premotor, temporal, and parietal cortex. *Neuroimage*, 22(1), 277-288.
- Wicker, B., Keysers, C., Plailly, J., Royet, J. P., Gallese, V., & Rizzolatti, G. (2003). Both of us disgusted in My insula: the common neural basis of seeing and feeling disgust. *Neuron*, 40(3), 655-664.
- Williams, M. A., Berberovic, N., & Mattingley, J. B. (2007). Abnormal fMRI adaptation to unfamiliar faces in a case of developmental prosopamnesia. *Curr Biol*, 17(14), 1259-1264.
- Yin, R. K. (1969). Looking at upside-down faces. *Journal of Experimental Psychology*, 81, 141-145.
- Zald, D. H. (2003). The human amygdala and the emotional evaluation of sensory stimuli. *Brain Res Brain Res Rev*, 41(1), 88-123.
- Zimmerman, E. K., Eslinger, P. J., Simmons, Z., & Barrett, A. M. (2007). Emotional perception deficits in amyotrophic lateral sclerosis. *Cogn Behav Neurol*, 20(2), 79-82.

Chapter 8: Prosopagnosia

We often take for granted the ease with which we recognize friends and foes. However, imagine that every time you encounter an acquainted face, it would seem no more familiar than the face of a total stranger. This most peculiar symptom is known as prosopagnosia and not surprisingly it puts a heavy burden on the social life of the patients who suffer from it.

Definitions and concepts

The typical complaint from a prosopagnosic regards the inability to recognize a person by the face. This symptom is far more pronounced than the phenomenon everybody sometimes experiences when have trouble remembering from where or how they know a certain face. Prosopagnosics can even have difficulties recognizing the persons they are very close to like their direct family members. Prosopagnosics have to rely on other cues to identify a person, for instance clothing or voice. But when they look at photographs taken in close-up, recognizing the person they are looking at becomes a challenge, even when they are looking at their own child. In less radical cases prosopagnosics are helped by the context in which they routinely encounter individuals but they fail to recognize them when outside the normal context, for example when they meet a colleague in the supermarket.

Prosopagnosia was first reported by Bodamer (1947). The deficit involves recognition of personal identity but not of facial expression and this dissociation has long been the cornerstone of the models of face processing in the neuropsychological literature of the last two decades and is at the basis of the face recognition model of Bruce & Young (1986). Until recently the dissociation between identity and expression was generally accepted and as a consequence researchers rarely investigated facial expression recognition itself nor was much attention paid to the possibility of an interaction between the face identity and the facial expression recognition system. When problems with faces are defined as problems with identity recognition, understanding face recognition may or may not be seen as a subprocess of object recognition. A substantial effort has been devoted to find evidence for dissociations between face and object recognition. In the eighties and nineties the gold standard for

supporting this kind of special or modular cognitive ability argument was to find a double dissociation. Patients with object recognition deficits that were not prosopagnosic vs. prosopagnosics without object recognition problems perfectly fitted this bill. Such pure cases have however proven to be extremely rare. Furthermore, the rich literature on reading difficulties provided a complementary window on the perceptual processing routes involved in face and object recognition. The most extensive review of available cases in the literature by Farah (1990) in the late eighties added a third category to the contrast. Patients with acquired dyslexia or inability to read words as a consequence of brain damage showed a pattern of letter by letter reading.

It is interesting to note that the neuropsychological approach focussed on these patient categories as a means to understand not so much category specificity per se, but the underlying perceptual routes that may or may not be common to face and word recognition deficits which both can be characterized as loss of the ability to process the configuration.

An older and often ignored effort at getting underneath the surface of a behavioural deficit was to understand prosopagnosia as a form of agnosia using a historical distinction between two types of agnosia, described by Lissauer (1890): apperceptive agnosia reflects impairment in consciously perceiving and recognizing stimuli and associative agnosia refers to a relatively preserved ability to perceive objects, but an inability to interpret what is seen. Apperceptive prosopagnosia therefore relates to an inability to perceive a face as a face and patients suffering from this type will be mostly impaired in tasks tapping into face *detection* and face *discrimination*, whereas associative prosopagnosia is in the first place associated with the recognition of individual faces and involves a deficit in the kind of encoding that allows telling individual exemplars apart and access stored representations and thereby also a memory component. It is the latter type that has been mainly reported and investigated as it is assumed that prosopagnosia arises at a higher object recognition and cognitive level. Questions on intact face detection typically are not raised about prosopagnosia and intact face detection is typically not seen as possibly a separate and earlier stage in face perception (but see B. de Gelder, Frissen, Barton, & Hadjikhani, 2003; B. de Gelder & Rouw, 2000a).

Inability to process the face as a whole

The most common functional characterisation of the core deficit of prosopagnosics is an impairment in processing the face at the level of its overall configuration or a deficit in configural processing. Yin (1969) showed that faces become much harder to recognize when they are presented upside down. This loss of performance for inverted presentation, named the inversion effect, has since provided researchers with a criterion for measuring the presence of configural processing skills. It is important to notice though that in the original study the inversion effect was defined not as unique to faces but as relatively more pronounced for faces. A number of other object categories tested, like for example landscapes were also recognized less well when presented upside down. The implication is that when using the inversion effect as a measure of intact configuration processing of faces, a task requirement needs to be used. Only when the inversion deficit is restricted to faces can one conclude to a face specific disorder at least as assessed by this effect (B. de Gelder & Rouw, 2000b).

Acquired and developmental prosopagnosia

Prosopagnosia can occur after a neurological antecedent, in which case it is known as acquired prosopagnosia (AP) and is most frequently associated with a lesion in a particular part of the visual cortex, namely the fusiform gyrus, especially in the right hemisphere (e.g. Barton, Press, Keenan, & O'Connor, 2002; B. de Gelder et al., 2003; Marotta, Genovese, & Behrmann, 2001), but other cases have been reported with more posterior lesions (Sorger, Goebel, Schiltz, & Rossion, 2007; Steeves et al., 2006) or following a head trauma but without evident brain lesions (B. de Gelder et al., 2003; Hadjikhani & de Gelder, 2002).

Prosopagnosia can also be present without any neurological history. In that case it is known as developmental prosopagnosia (DP). Developmental prosopagnosics usually report life long problems in recognizing people by the face. Developmental prosopagnosia sometimes occurs in different generations of a single family, suggesting a genetic component, more particularly an autosomal dominant mode of inheritance (Grueter et al., 2007). The term 'congenital prosopagnosia' is sometimes used to refer to prosopagnosia without any neurological antecedents. However, the genetic basis of DP is a matter of debate and there is no evidence

to conclude that all these cases have a genetic basis. The term ‘developmental prosopagnosia’ has a more descriptive connotation and includes, next to the genetic hypothesis, the possibility of a developmental disorder, for instance in the maturation of different cortical areas, like the fusiform gyrus. A study investigating the volumetric and morphological properties of DP brains has revealed a decreased size of the fusiform gyrus and this decrease correlated with the behavioral face recognition difficulties of the subjects (Behrmann, Avidan, Gao, & Black, 2007). The prevalence of DP has been estimated at 2.47% (Kennerknecht, Plumpe, Edwards, & Raman, 2006).

Face-specificity of prosopagnosia

The face recognition deficit rarely occurs in isolation, but is usually accompanied by problems in recognizing other stimuli like cars or places. Faces provide many kinds of information like gender, age and emotional expression and a finely tuned comparison of face recognition abilities with other object recognition abilities calls for comparable task settings between the two object classes (Damasio, Damasio, & Van Hoesen, 1982; Damasio, Tranel, & Damasio, 1990; B. de Gelder, Bachoud-Levi, & Degos, 1998; Farah, 1990; Gauthier, Behrmann, & Tarr, 1999). This is a challenging task, since faces have many dimensional attributes and are all around us from the moment we are born, which means the perceptual system has been frequently exposed to them. Consequently, this has led to an ongoing debate about whether the perception of faces is qualitatively different from the perception of objects (modularity hypothesis) (Fodor, 1983), or on the other hand that the ‘special’ status of face processing merely reflects the level of perceptual expertise we have with them (expertise hypothesis) (Diamond & Carey, 1986; Gauthier, Skudlarski, Gore, & Anderson, 2000).

One object category that has comparable properties to faces is human bodies, which also carry information about age, gender, emotional expression, etc (for reviews, see B. de Gelder, 2006; Peelen & Downing, 2007). A behavioural study with a single prosopagnosic case who was instructed to match the identity of computer generated bodies, found no evidence for impaired body recognition (B. C. Duchaine, Yovel, Butterworth, & Nakayama, 2006), while an event-related potential (ERP) study with four developmental prosopagnosics (DPs) found abnormal brain activation in the early time window of the EEG for both faces and bodies in three of the four DPs (Righart & de Gelder, 2007). Only one study investigated

the functional neuro-anatomy of perceiving faces and bodies in three DPs and matched control subjects. The results showed that perceiving bodies activates face areas in the brains of DPs and also that perceiving faces activates body areas (Van den Stock, van de Riet, Righart, & de Gelder, 2008). This shows that the functional neuro-anatomy of perceiving either faces or bodies shows reduced specialization in DP.

Although face recognition difficulties are the core feature of prosopagnosia, the visual problems are not always confined to the category of faces and patients also report difficulties in recognizing other object classes like houses or cars (e.g. Bobes et al., 2003). Although the term 'prosopagnosia' strictly does not exclude object recognition difficulties, standard procedures in prosopagnosia assessment must include at least a basic screening of the visual object recognition abilities with standard clinical test batteries like the Birmingham Object Recognition Battery (Riddoch & Humphreys, 1993) or the Visual Object and Space Perception Battery (E.K. Warrington & James, 1991). It is not exceptional that prosopagnosics score outside of the normal range on at least one of the tests (e.g. B. C. Duchaine, 2000; Righart & de Gelder, 2007). Standard clinical tests evaluating face matching and memory for faces include the Benton Facial Recognition Test (BFRT) (Benton, Hamsher, Varney, & Spreen, 1983) and the Warrington Recognition Memory Test (WRMT) (Warrington, 1984). While some prosopagnosics obtain deficient scores on these tests of face processing (e.g. Ariel & Sadeh, 1996; B. de Gelder & Rouw, 2000a; Righart & de Gelder, 2007), others perform within the normal range (e.g. B. C. Duchaine, 2000; Nunn, Postma, & Pearson, 2001), hence the scores on these tests should be cautiously interpreted, especially when no reaction time data is available as is the case in the older studies.

More modern tests and batteries are now available targeting specific aspects of face and object processing (e.g. B. de Gelder et al., 1998; B. de Gelder & Rouw, 2000a, 2000b; B. de Gelder & Van den Stock, 2008; B. Duchaine & Nakayama, 2006). These provide a more detailed analysis of the intact and anomalous aspects of face and object processing. The Facial Expressive Action Stimulus Test (FEAST) (B. de Gelder & Van den Stock, 2008) for example targets a range of specific abilities like face detection, emotional face memory, neutral face memory, facial emotion matching, face identity matching, object identity matching, face-part matching, configural face processing. The latter refers to the notion that faces are processed as a whole, whereas other objects are more processed as an assemblage of features. Configural processing is classically measured by the inversion effect: upside-down presented faces result

in a (larger) decline in performance (than upside down presented objects). It has been shown that some DPs display the inverse pattern, namely a paradoxical inversion effect: they perform better when faces are presented upside down, but also when objects are presented upside down (B. de Gelder & Rouw, 2000b).

Behavioural and neural correlates of prosopagnosia

Functional Magnetic Resonance Imaging (fMRI) studies investigating the neural correlates of DP have yielded inconsistent results (Avidan, Hasson, Malach, & Behrmann, 2005; Bentin, Degutis, D'Esposito, & Robertson, 2007; Hadjikhani & de Gelder, 2002; Hasson, Avidan, Deouell, Bentin, & Malach, 2003). In neurologically intact subjects, perception of faces selectively activates a network of brain areas in the right occipito-temporal cortex, namely the fusiform gyrus (FG): the so-called fusiform face area (FFA) (Haxby et al., 1994; Kanwisher, McDermott, & Chun, 1997; Sergent, Ohta, & MacDonald, 1992) and the inferior occipital gyrus (IOG) (Gauthier et al., 2000; Hoffman & Haxby, 2000; Puce, Allison, Asgari, Gore, & McCarthy, 1996). In the first fMRI-study with a DP patient, these areas did not show the typical activation pattern as observed in normals (2002) when contrasting faces with objects. Similar results were obtained in another study with a DP (Bentin et al., 2007). A normal activation of these areas was found in two studies with respectively one and four DPs (Avidan et al., 2005; Hasson et al., 2003). However, only one study investigated the neural correlates of processing both neutral and emotional faces and found that there was no difference in FFA activation between controls and three DPs, but only when these faces express an emotion (fear or happiness). When perceiving neutral faces, the FFA of DPs showed a decline in activation, compared to the control subjects (Van den Stock et al., 2008).

Since normal face-processing involves multiple hierarchical and parallel processes, impairments in different processes will result in different types of behavioral and neuro-anatomical correlates. In developmental disorders like DP, heterogeneity is rather the rule than the exception and this is also reflected in the behavioral findings. Extensive evaluations of single cases or small groups of prosopagnosics have found different patterns of impairment on multiple aspects of face processing, including configural processing, modulation of the N170, within-object spatial relations, recognition of emotions, gender discrimination, recognition of famous faces, holistic face processing and facial attractiveness ratings.

References

- Ariel, R., & Sadeh, M. (1996). Congenital visual agnosia and prosopagnosia in a child: a case report. *Cortex*, 32(2), 221-240.
- Avidan, G., Hasson, U., Malach, R., & Behrmann, M. (2005). Detailed exploration of face-related processing in congenital prosopagnosia: 2. Functional neuroimaging findings. *J Cogn Neurosci*, 17(7), 1150-1167.
- Barton, J. J., Press, D. Z., Keenan, J. P., & O'Connor, M. (2002). Lesions of the fusiform face area impair perception of facial configuration in prosopagnosia. *Neurology*, 58(1), 71-78.
- Behrmann, M., Avidan, G., Gao, F., & Black, S. (2007). Structural imaging reveals anatomical alterations in inferotemporal cortex in congenital prosopagnosia. *Cereb Cortex*, 17(10), 2354-2363.
- Bentin, S., Degutis, J. M., D'Esposito, M., & Robertson, L. C. (2007). Too Many Trees to See the Forest: Performance, Event-related Potential, and Functional Magnetic Resonance Imaging Manifestations of Integrative Congenital Prosopagnosia. *J Cogn Neurosci*, 19(1), 132-146.
- Benton, A. L., Hamsher, K., Varney, N. R., & Spreen, O. (1983). *Test of facial recognition: Form SL*. New York: Oxford University Press.
- Bobes, M. A., Lopera, F., Garcia, M., Diaz-Comas, L., Galan, L., & Valdes-Sosa, M. (2003). Covert matching of unfamiliar faces in a case of prosopagnosia: an ERP study. *Cortex*, 39(1), 41-56.
- Bodamer, J. (1947). Die prosop-Agnosie. *Archiv fur Psychiatrie und Nervenkrankheiten*, 179, 6-53.
- Bruce, V., & Young, A. (1986). Understanding face recognition. *Br J Psychol*, 77 (Pt 3), 305-327.
- Damasio, A. R., Damasio, H., & Van Hoesen, G. W. (1982). Prosopagnosia: anatomic basis and behavioral mechanisms. *Neurology*, 32(4), 331-341.
- Damasio, A. R., Tranel, D., & Damasio, H. (1990). Face agnosia and the neural substrates of memory. *Annu Rev Neurosci*, 13, 89-109.
- de Gelder, B. (2006). Towards the neurobiology of emotional body language. *Nature Reviews Neuroscience*, 7(3), 242-249.

- de Gelder, B., Bachoud-Levi, A. C., & Degos, J. D. (1998). Inversion superiority in visual agnosia may be common to a variety of orientation polarised objects besides faces. *Vision Res*, 38(18), 2855-2861.
- de Gelder, B., Frissen, I., Barton, J., & Hadjikhani, N. (2003). A modulatory role for facial expressions in prosopagnosia. *Proc Natl Acad Sci U S A*.
- de Gelder, B., & Rouw, R. (2000a). Configural face processes in acquired and developmental prosopagnosia: evidence for two separate face systems? *Neuroreport*, 11(14), 3145-3150.
- de Gelder, B., & Rouw, R. (2000b). Paradoxical configuration effects for faces and objects in prosopagnosia. *Neuropsychologia*, 38(9), 1271-1279.
- de Gelder, B., & Van den Stock, J. (2008). Facial Expressive Action Stimulus Test (FEAST).
- Diamond, R., & Carey, S. (1986). Why faces are and are not special: an effect of expertise. *J Exp Psychol Gen*, 115(2), 107-117.
- Duchaine, B., & Nakayama, K. (2006). The Cambridge Face Memory Test: results for neurologically intact individuals and an investigation of its validity using inverted face stimuli and prosopagnosic participants. *Neuropsychologia*, 44(4), 576-585.
- Duchaine, B. C. (2000). Developmental prosopagnosia with normal configural processing. *Neuroreport*, 11(1), 79-83.
- Duchaine, B. C., Yovel, G., Butterworth, E. J., & Nakayama, K. (2006). Prosopagnosia as an impairment to face-specific mechanisms: Elimination of the alternative hypotheses in a developmental case. *Cognitive Neuropsychology*, 23(5), 714-747.
- Farah, M. (1990). *Visual agnosia: Disorders of visual recognition and what they tell us about normal vision*. Cambridge: MIT Press.
- Fodor, J. (1983). *The Modularity of Mind*. Cambridge, MA: MIT Press.
- Gauthier, I., Behrmann, M., & Tarr, M. J. (1999). Can face recognition really be dissociated from object recognition? *J Cogn Neurosci*, 11, 349-370.
- Gauthier, I., Skudlarski, P., Gore, J. C., & Anderson, A. W. (2000). Expertise for cars and birds recruits brain areas involved in face recognition. *Nature Neuroscience*, 3(2), 191-197.
- Grueter, M., Grueter, T., Bell, V., Horst, J., Laskowski, W., Sperling, K., et al. (2007). Hereditary prosopagnosia: the first case series. *Cortex*, 43(6), 734-749.

- Hadjikhani, N., & de Gelder, B. (2002). Neural basis of prosopagnosia: an fMRI study. *Hum Brain Mapp*, 16(3), 176-182.
- Hasson, U., Avidan, G., Deouell, L. Y., Bentin, S., & Malach, R. (2003). Face-selective activation in a congenital prosopagnosic subject. *Journal of Cognitive Neuroscience*, 15(3), 419-431.
- Haxby, J. V., Horwitz, B., Ungerleider, L. G., Maisog, J. M., Pietrini, P., & Grady, C. L. (1994). The functional organization of human extrastriate cortex: a PET-rCBF study of selective attention to faces and locations. *J Neurosci*, 14(11 Pt 1), 6336-6353.
- Hoffman, E. A., & Haxby, J. V. (2000). Distinct representations of eye gaze and identity in the distributed human neural system for face perception. *Nat Neurosci*, 3(1), 80-84.
- Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: a module in human extrastriate cortex specialized for face perception. *J Neurosci*, 17(11), 4302-4311.
- Kennerknecht, I., Plumpe, N., Edwards, S., & Raman, R. (2006). Hereditary prosopagnosia (HPA): the first report outside the Caucasian population. *J Hum Genet*.
- Lissauer, H. (1890). Ein Fall von Seelenblindheit nebst einem Beitrage zur Theorie derselben.[A case of visual agnosia with a contribution to theory.]. *Archiv für Psychiatrie und Nervenkrankheiten*, 21, 222-270.
- Marotta, J. J., Genovese, C. R., & Behrmann, M. (2001). A functional MRI study of face recognition in patients with prosopagnosia. *Neuroreport*, 12(8), 1581-1587.
- Nunn, J. A., Postma, P., & Pearson, R. (2001). Developmental prosopagnosia: should it be taken at face value? *Neurocase*, 7(1), 15-27.
- Peelen, M. V., & Downing, P. E. (2007). The neural basis of visual body perception. *Nat Rev Neurosci*, 8(8), 636-648.
- Puce, A., Allison, T., Asgari, M., Gore, J. C., & McCarthy, G. (1996). Differential sensitivity of human visual cortex to faces, letterstrings, and textures: a functional magnetic resonance imaging study. *J Neurosci*, 16(16), 5205-5215.
- Riddoch, M. J., & Humphreys, G. W. (1993). *Birmingham Object Recognition Battery*. Hove: Psychology Press.
- Righart, R., & de Gelder, B. (2007). Impaired face and body perception in developmental prosopagnosia. *Proc Natl Acad Sci U S A*.

- Sergent, J., Ohta, S., & MacDonald, B. (1992). Functional neuroanatomy of face and object processing. A positron emission tomography study. *Brain*, 115(Pt 1), 15-36.
- Sorger, B., Goebel, R., Schiltz, C., & Rossion, B. (2007). Understanding the functional neuroanatomy of acquired prosopagnosia. *Neuroimage*, 35(2), 836-852.
- Steeves, J. K., Culham, J. C., Duchaine, B. C., Pratesi, C. C., Valyear, K. F., Schindler, I., et al. (2006). The fusiform face area is not sufficient for face recognition: evidence from a patient with dense prosopagnosia and no occipital face area. *Neuropsychologia*, 44(4), 594-609.
- Van den Stock, J., van de Riet, W. A., Righart, R., & de Gelder, B. (2008). Neural correlates of perceiving emotional faces and bodies in developmental prosopagnosia: an event-related fMRI-study. *PLoS ONE*, 3(9), e3195.
- Warrington, E. K. (1984). *Recognition Memory Test*. Nelson, Windsor: NFER.
- Warrington, E. K., & James, M. (1991). *Visual Object and Space Perception Battery (VOSP)*. Bury St. Edmunds, UK: Thames Valley Test Company.
- Yin, R. K. (1969). Looking at upside-down faces. *Journal of Experimental Psychology*, 81, 141-145.

Chapter 9: Neural correlates of perceiving emotional faces and bodies in developmental prosopagnosia: an event-related fMRI-study⁸

Introduction

Recognizing faces of family and friends usually proceeds effortlessly. Yet a minority of people has difficulties telling apart who they are meeting with or remembering who they met previously when they can only go by the visual memory of the face. These problems can be quite dramatic, even to the point where they fail to recognize the face of their own spouse or child or for that matter their own face. The original reports of face recognition deficits for which the term prosopagnosia (Bodamer, 1947) was coined concerned cases of brain damage sustained in adulthood. More recently there have been reports of face recognition deficits that do not appear to be associated with any known neurological history. Although there are still only a few systematic reports of this condition, many more cases are described now compared to a decade ago and some authors have argued that as much as 2% of the population suffers from face recognition difficulties (Kennerknecht et al., 2006). In analogy with developmental dyslexia these cases are now commonly referred to as developmental prosopagnosia (DP), referring to the possible origin of the adult face recognition deficit in anomalous development of the full face recognition skills. This behavioral deficit may include an anomaly in the putative congenital basis involved in the acquisition of the skill, but so far very little is known about this genetic basis and its importance for explaining behavioral deficits (Grueter et al., 2007).

Recent research on behavioral face recognition deficits and their neural basis has followed the leads from the reports on the neural basis of face recognition in normals as mainly revealed in fMRI studies over the last decade. There is now a consensus in the literature that face recognition is implemented in a network of brain areas (B. de Gelder & R. Rouw, 2000a; Haxby, Hoffman, & Gobbini, 2000). Among these, an area in the fusiform gyrus (FG), labeled the fusiform face area (FFA) (Grill-Spector, Knouf, & Kanwisher, 2004; Kanwisher, McDermott, & Chun, 1997), has attracted most attention. Next to this area, the

⁸ This chapter is based on: Van den Stock, J., van de Riet, W.A.C., Righart, R., de Gelder, B. (2008). Neural correlates of perceiving emotional faces and bodies in developmental prosopagnosia: an event-related fMRI-study. *PLoS ONE*, 3(9): e3195.

role of the inferior occipital gyrus (IOG) is repeatedly stressed in normal (e.g. Gauthier, Skudlarski, Gore, & Anderson, 2000; Hoffman & Haxby, 2000; Puce, Allison, Asgari, Gore, & McCarthy, 1996) and anomalous face recognition (Hadjikhani & de Gelder, 2002). But it is fair to say that the functional significance of these two main areas for person recognition and its deficits is not yet entirely clear.

Investigations of the neuro-functional correlates of DP with fMRI have yielded inconsistent results (Avidan, Hasson, Malach, & Behrmann, 2005; Bentin, Degutis, D'Esposito, & Robertson, 2007; Degutis, Bentin, Robertson, & D'Esposito, 2007; Hadjikhani & de Gelder, 2002; Hasson, Avidan, Deouell, Bentin, & Malach, 2003; Williams, Berberovic, & Mattingley, 2007) (see Table 1 for an overview). The first fMRI-study including a DP case by Hadjikhani and de Gelder (2002) found no face-specific activation in these two areas. A similar pattern was observed with another DP case (Bentin et al., 2007). On the other hand, other studies reported normal face-specific activation in developmental prosopagnosics (DPs) despite their severe behavioral deficits in face recognition (Avidan et al., 2005; Degutis et al., 2007; Hasson et al., 2003; Williams et al., 2007). These findings suggest that intact functioning of the FFA and IOG are necessary, but not sufficient for successful face recognition.

Table 1: Results from fMRI-studies on prosopagnosia. Abbreviations: n.a.: not applicable; FFA: Fusiform Face Area; FG: Fusiform Gyrus; IOG: Inferior Occipital Gyrus; SM, CR, GA & RP refer to subjects; +: significant activation; -: no significant activation; (l-): only left activation is observed

Studies with developmental prosopagnosics						
	N	Lesion localisation	task	comparison	result	
					FFA	IOG
Hadjikhani & de Gelder (2002)	1	n.a.	Passive viewing	Faces>objects	-	-
				Faces>houses	+	-
Hasson et al. (2003)	1	n.a.	One back	Faces>buildings	+	+
Avidan et al. (2005)	4	n.a.	One back	Faces>(buildings & objects)	+	+
Bentin et al. (2007)	1	n.a.	Oddball	Faces>places	-	-

				Faces>objects	-	-
Degutis et al. (2007) ⁹	1	n.a.	One back	Faces>scenes	+	+
Williams et al. 2007	1	n.a.	One back	Faces>scenes	+	
Van den Stock et al.	3	n.a.	Oddball	(Emotional + neutral faces)>houses	decreased neutral face activity	+
Studies with acquired prosopagnosics						
	N	Lesion localisation	task	comparison	result	
					FFA	IOG
Marotta et al. (2001)	2	SM: right temporal	oddball	Faces>objects	SM	(+)
		CR: right temporal			CR	-
Hadjikhani & de Gelder (2002)	2	No evident lesions	Passive viewing	Faces>objects	-	-
				GA: Faces>scramble	-	-
				RP: Faces>houses	-	+
Rossion et al. (2003)	1	Right posterior inferior occipital to posterior fusiform gyrus; left middle fusiform gyrus	One back	Faces>objects	+	
Steeves et al. (2006)	1	Bilateral lateral occipital; left superior parieto-occipital sulcus	Oddball	Faces>scenes	+	
Sorger et al. (2007)	1	Right ventral occipito-temporal; right middle temporal gyrus; left FG, IOG and lingual gyrus; left medial cerebellum (see also Rossion et al. 2003)	One back	Faces>objects	+	(l)

In view of the many different kinds of information a face provides (gender, age, emotion, familiarity, attractiveness etc.) and the different ways in which this information is called upon and used in daily life (whether the context only requires rapid detection that there is a

⁹ This study reports about a training program administered to a patient. We report the fMRI result preceding the training.

face present, or on the contrary, full recognition of all facial attributes including name retrieval), it is worth stressing that the contextual requirements and the task settings are very important for evaluating face recognition problems and for understanding its neuro-functional basis and possible deficits. A finely tuned comparison of face recognition skills with other object recognition skills at the behavioral and neuro-functional level requires comparable task settings whether the object categories to be matched are faces or any other category that is suitable (Damasio, Damasio, & Van Hoesen, 1982; Damasio, Tranel, & Damasio, 1990; de Gelder, Bachoud-Levi, & Degos, 1998; Farah, 1990; Gauthier, Behrmann, & Tarr, 1999). Since faces convey many different kinds of information it has so far been a daunting task to find a matching category to use as control stimuli. Previous approaches to find the best matching category have tended to explore either the physical similarity dimension (for example, using a continuum of more or less face like stimuli), the perceptual one or the functional one (for example, expertise with one or another specific object category). This has fed an ongoing debate about whether face processing mechanisms are qualitatively different from the processing mechanisms for objects (modularity hypothesis) (Fodor, 1983), or on the other hand whether relative face specificity reflects the level of perceptual expertise with the stimulus category (expertise hypothesis) (Diamond & Carey, 1986; Gauthier et al., 2000). As a matter of fact there are very few objects other than faces for which strong claims about category specific representation have been made. One exception concerns houses. Several studies report that this object category differentially activates a region around the collateral sulcus (Aguirre, Zarahn, & D'Esposito, 1998; Epstein & Kanwisher, 1998; Levy, Hasson, Avidan, Hendler, & Malach, 2001).

An interesting object category not used so far concerns human bodies. Recently, it has been shown in normal subjects that perceiving human bodies or body parts activates an area in extrastriate cortex, labeled extrastriate body area (EBA) (Downing, Jiang, Shuman, & Kanwisher, 2001). More recently a second body specific area was defined in the FG (Hadjikhani & de Gelder, 2003; Peelen & Downing, 2005). This body sensitive area in FG overlaps at least partially with the face-sensitive one and it has been termed the fusiform body area (FBA). In parallel, recent findings show that the close similarities between face and body perception exist at the level of perceptual mechanisms as revealed by the inversion effect (a decline in performance for inverted stimuli compared to upright stimuli that is more pronounced for faces than for other object categories (Yin, 1969)), since the same inversion

effect has been reported for bodies (Reed, Stone, Bozova, & Tanaka, 2003; Stekelenburg & de Gelder, 2004) (for reviews, see de Gelder, 2006; Peelen & Downing, 2007).

These behavioral and neuro-functional similarities between perceiving faces and bodies in normals and the fact that bodies represent a distinct but yet very closely related object category, raise the issue how bodies are processed in DP. A study by Duchaine et al. (2006) presented natural faces and computer generated neutral body postures for testing face and body identity recognition in a DP patient using a sequential identity matching paradigm involving a minimal memory component. The performance of the patient was impaired for the faces, but within normal range for the bodies suggesting dissociation between face and body processing mechanisms with these task settings. Another study used event-related potentials (ERP) to investigate face and body perception in four DPs and found abnormal brain activation in the early time windows of the EEG (around 170 ms) for both faces and bodies in three of the four DPs (Righart & de Gelder, 2007).

A second main objective of the present study is to investigate how the neural underpinnings of face and body processing in prosopagnosia are influenced by emotional information in the face and the body. As a matter of fact, the face-sensitive area in FG is well known from investigations of face recognition using neutral faces but it also figures predominantly in research on the neural basis of recognizing facial expressions. The presence of an emotion expression adds realism to the face but may also be an interesting developmental factor. Studies with younger subjects have predominantly reported higher activation for fearful faces, compared to neutral faces (Breiter et al., 1996; Dolan et al., 1996; Dolan, Morris, & de Gelder, 2001; Rotshtein, Malach, Hadar, Graif, & Hendler, 2001; Vuilleumier, Armony, Driver, & Dolan, 2001), but a recent study with both adolescents and adults found a reverse pattern in the FFA, namely higher activation for neutral than for fearful faces (Guyer et al., 2008). The mechanism of this emotional modulation in the FFA may be based on feedback loops with the amygdala (Breiter et al., 1996; Vuilleumier et al., 2001). A similar explanation has been proposed for the increased activation in FG sensitive to body images representing an emotional expression (Hadjikhani & de Gelder, 2003).

So far, the evidence concerning the neural correlates of processing emotional faces in DP is scarce. One study by de Gelder et al. (2003) investigated this issue in acquired prosopagnosics (prosopagnosia occurring after brain damage). The included patients had

lesions in either the FG, IOG or both. The results showed that the patients more strongly activated other face sensitive areas like the superior temporal sulcus (STS) or amygdala when they perceive facial expressions compared to neutral faces. The patients were also more accurate and faster in processing emotional faces compared to neutral faces, a finding that has been reported previously (Duchaine, Parker, & Nakayama, 2003; Jones & Tranel, 2001; Nunn, Postma, & Pearson, 2001). Since the patients in de Gelder et al. (2003) had lesions in the ventral occipito-temporal cortex, the question arises how these brain areas respond to emotional information in prosopagnosics with severe face recognition problems but no known brain anomalies. To investigate this issue we presented the participants with neutral, fearful and happy facial and bodily expressions.

Methods

Participants. The DPs were recruited after they had contacted us via our website or through reports in the popular press. All participants report life-long problems in recognizing people and typically complain about difficulties when meeting familiar persons unexpectedly and the ensuing social problems. AM (female) is a 54-year old housewife. She reports problems in recognizing others when meeting them outside the usual context, for example when she meets her parents in the supermarket. HV (male) is 43 years old and teaches writing and coaches in communication training. He experiences severe face recognition problems for as long as he can remember. LW (male) is a 48-year old university professor with longstanding difficulties for example in recognizing colleagues at conferences and students. None of the DPs had a neurological history and their structural MR-scans showed no abnormalities as judged independently by four experienced neurologists. The group of four control subjects was matched with the DP group on age, sex and educational level. All participants gave written informed consent according to the Declaration of Helsinki and the study was approved by the local ethics committee (CMO region Arnhem-Nijmegen, The Netherlands).

Neuropsychological testing

All participants were presented with an extensive face recognition battery. Visual object recognition and face recognition were assessed with standard clinical tests and additional face and object perception experiments were run in sessions preceding the fMRI measurements. The neuropsychological tests and normative data are described elsewhere (Righart & de Gelder, 2007). Face matching and face memory were tested with the Benton Face Recognition Test (BFRT) (Benton, Sivan, Hamsher, Varney, & Spreen, 1983) and the Warrington Face Memory Test (WFMT) (Warrington, 1984). We used a computerized version of the latter test to obtain information about speed-accuracy trade-off. Basic visual functions were measured with the Birmingham Object Recognition Battery (BORB) (line length, size, orientation, gap, minimal feature match, foreshortened views and object decision) (Riddoch & Humphreys, 1993). To investigate in detail different aspects of face perception, all participants were administered additional face and object perception experiments which have proven useful in previous investigations of face recognition and provided insight in processing strategies in prosopagnosia (de Gelder et al., 1998; de Gelder et al., 2003; B. de Gelder & P. Rouw, 2000; B. de Gelder & R. Rouw, 2000a, 2000b; Righart & de Gelder, 2007).

Like in our previous studies on prosopagnosia, the behavioral pattern of a normal inversion effects for faces compared to another single object category was measured with the faces and shoes task (de Gelder et al., 1998). Participants were required to select the probe that corresponded with the identity of a simultaneously presented target. The target was always a frontal picture and the two probes underneath consisted of pictures in three quarter profile. Faces and shoes were presented upright and inverted (for details, see de Gelder et al., 1998; B. de Gelder & R. Rouw, 2000b). Feature-based processing was tested with a part-to-whole matching task which required participants to select the face-part probe (i.e., mouth or eyes) that was the same as that in the simultaneously presented whole face. The same procedure was followed for house-part probes (i.e., door or upper window) that had to be matched to the corresponding part in a whole house stimulus. Faces and houses were presented once upright and once inverted (de Gelder et al., 2003; B. de Gelder & R. Rouw, 2000a). Participants were instructed to respond as accurately and rapidly as possible. Accuracy and mean response-times were calculated for each test. We compared the accuracy

and response times from the upright stimuli with the inverted stimuli in one-tailed paired-sample t-tests. A significantly lower accuracy or longer response time for the inverted stimuli is defined as an inversion effect, whereas a higher accuracy or shorter response time for the inverted stimuli is defined as a paradoxical inversion effect. Data of the control group were normalized and z-scores were obtained for every DP.

fMRI measurements

Stimulus materials

The face and body stimuli were used previously in an fMRI investigation of the neural substrates of processing face and body perception in neurologically intact observers (van de Riet, Grèzes, & de Gelder, in press). Pictures of fearful, happy and neutral faces were taken from the Karolinska Directed Emotional Face database (Lundqvist, Flykt, & Öhman, 1998). From our own database, pictures of fearful and happy bodily expressions, instrumental (emotionally neutral) bodily expressions (pouring water into a glass) and houses were used. We used houses as stimuli for the control condition, because they constitute a single object category that has been extensively explored in other studies and is known to elicit activation in specific brain areas (Aguirre et al., 1998; Epstein & Kanwisher, 1998; Levy et al., 2001). Instrumental body expressions were used because, like emotional expressions, these displays elicit action representation and implicit movement (Johnson-Frey et al., 2003), and hence constitute a balanced comparison category for the emotional expressions. All images of faces and bodies were previously validated regarding emotional expression (minimum recognition rate: 75%). (For further details concerning the validation procedure, see van de Riet et al., in press).

A total of 42 images was used, six in every condition (fearful faces, happy faces, neutral faces, fearful bodies, happy bodies, neutral bodies and houses). There was no identity overlap between faces and bodies or between the emotions. Faces were fitted inside a gray oval shape, which masked external aspects of the faces. Body and house stimuli were cut out, removing all background. The faces of the body stimuli were covered with a gray opaque mask.

Additionally, one picture of a chair was used as an oddball stimulus. All stimuli were resized to 300 pixels in height and presented on a gray background.

Procedure

The design was adapted from our previous study (in press). In order not to exacerbate the face handicap of the DP group, we modified the experimental paradigm from a facial expression categorization task to an oddball detection task thereby also avoiding selective attention to the faces with an emotional expression. Moreover, this procedure excludes that activation profiles are contaminated by motor responses in the conditions of interest while still providing control data on attention to the stimuli. A trial started with the presentation of a fixation cross (200ms), followed by a stimulus (500ms) and finally by a gray screen (2200ms) (see Figure 1). All stimuli were presented six times in random order in an oddball paradigm (participants were instructed to press a response button when a chair was shown). The session consisted of 288 trials (7 conditions x 6 identities x 6 presentations, plus 36 oddball trials). Additionally, 96 null-events consisting of a gray screen lasting the whole trial length were included to reduce stimulus onset predictability and to establish a baseline (Friston, Zarahn, Josephs, Henson, & Dale, 1999). The experiment was preceded by a short practice-session which used a different set of face and body stimuli.

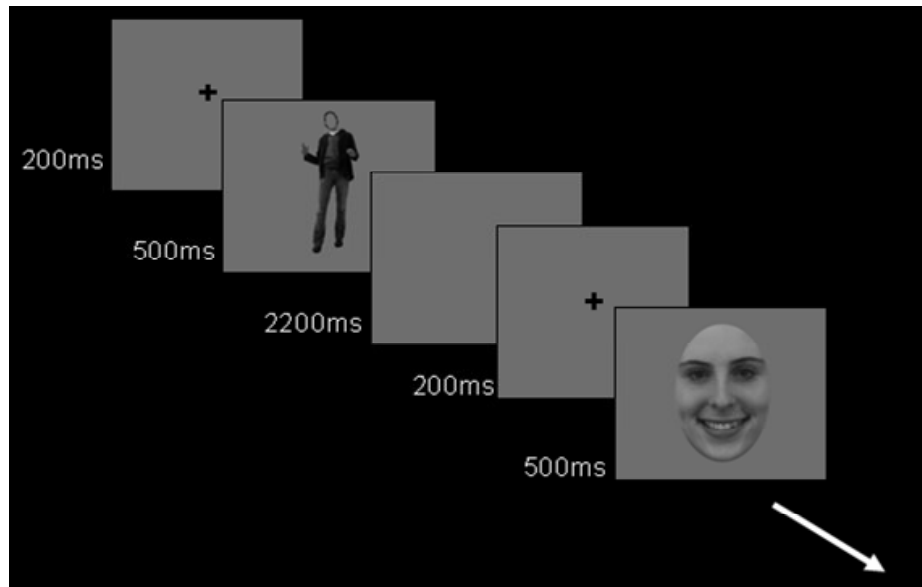


Figure 1. Schematic representation of the experimental design. Participants were instructed to press the response button when a chair was presented.

Participants lay supine in the scanner with head movements minimized by an adjustable padded head holder. Stimuli were projected onto a mirror above the participant's head. Responses were recorded via an MR-compatible keypad (MRI Devices, Waukesha, WI), positioned on the right side of the participant's abdomen. A PC running Presentation 9.70 (Neurobehavioral Systems, San Francisco, CA) controlled stimulus presentation and response registration.

Image Acquisition

Images were acquired using a 1.5 Tesla Sonata scanner (Siemens, Erlangen, Germany). Blood oxygenation level depend (BOLD) sensitive functional images were acquired using a single shot gradient echo-planar imaging (EPI) sequence [TR (repetition time)= 3790ms, TE (echo time)=40 ms, 43 transversal slices, ascending acquisition, 2.5mm slice thickness, with 0.25mm gap, FP (flip angle)=90°, FOV (field of view)= 32cm]. An automatic shimming procedure was performed before each scanning session. A total of 312 functional volumes were collected for each participant. Following the experimental session,

structural images were acquired using an MP-RAGE sequence [TR/TE/TI (inversion time) 2250 ms/3.93 ms/850 ms, voxel size $1 \times 1 \times 1$ mm].

Results

Neuropsychological testing

All DPs scored outside the normal range for the BFRT and/or the WFMT, but none showed an anomalous score on more than one subtest of the BORB suggesting that the visual recognition difficulties of the DPs as measured by these two clinical tests are not due to basic visual perception problems diagnosed in the BORB (see Table 2). AM scored significantly below the mean on the BFRT and WFMT, for both accuracies and response times. HV had a borderline performance on the BFRT and prolonged response times on the WFMT. LW scored within normal range on the BFRT, but on the WFMT both accuracy and response times were anomalous.

Table 2: Results from neuropsychological testing. Response times are shown for correct responses. Comparisons of DP's and matched controls are made by z-scores on the basis of the following control groups: Control group for the Warrington face memory: N = 25 (18-27 yrs); Control group for the Faces and Shoes task: N = 11 (18-28 yrs); Control group for the Face- and Houseparts: N = 21 (18-29 yrs); Asterisks indicate P-values corresponding to the Z-scores. * $p < .05$; ** $p < .01$; *** $p < .001$. SI: severe impairment; BL: borderline; A: average

		Controls	AM	HV	LW
BFRT accuracy (/54)		45.4 (A)	28 (SI)	40 (BL)	44 (A)
WFM accuracy (/50)		44.0	29***	41	34**
WFM RT (ms)		1778	3171***	3853 ***	3171 ***
Faces accuracy (/64)	Upr	63.3	57***	63	64
	Inv	62.0	56**	62	57**
Shoes accuracy (/64)	Upr	62.5	64	64	61
	Inv	62.8	62	64	58**
Faces RT (ms)	Upr	1146	3743 ***	2840 ***	1741**

	Inv	1526	3406 ***	3640 ***	2112
Shoes RT (ms)	Upr	978	2533 ***	1757 ***	1450***
	Inv	1069	2849 ***	1689 ***	1673**
Face-parts accuracy (/64)	Upr	62.7	47***	60*	59**
	Inv	62.0	52***	63	56**
House-parts accuracy (/64)	Upr	62.7	63	62	61
	Inv	63.2	64	64	63
Faceparts RT (ms)	Upr	1562	2099**	4446 ***	3462***
	inv	1755	2224	4130 ***	3229***
Houseparts RT (ms)	Upr	1192	1554*	1703 ***	1917***
	inv	1132	1361	1593 ***	1774***

To measure face and object recognition in a comparable way and assess relative configural processing routines, we compared upright and inverted stimulus matching for each object category (de Gelder et al., 1998; Righart & de Gelder, 2007). The control group showed an inversion effect for matching faces in both the accuracy ($t(10) = 1.892$, $p < .05$) and response time ($t(10) = 3.164$, $p < .005$). The controls showed no inversion effect for matching shoes. For the DPs, the response times were high as previously reported (B. de Gelder & R. Rouw, 2000a; Righart & de Gelder, 2007). AM was impaired in matching both upright ($Z < -5.75$) and inverted ($Z < -3.39$) faces. Her response times showed a paradoxical inversion effect pattern for matching faces and a normal inversion for matching shoes. HV had accuracies within the normal range, but displayed a normal inversion pattern in the response times for matching faces and a paradoxical inversion effect in the response times for matching shoes. LW showed reduced accuracy for matching inverted faces ($Z < -2.82$) and inverted shoes ($Z < -2.74$). His response times for matching upright faces were prolonged ($Z > 2.39$), while the latencies for inverted faces were on average. He displayed the normal inversion pattern for matching faces and shoes in both accuracy and response times.

Feature-based matching was tested with the faces and houses task (see B. de Gelder & R. Rouw, 2000a for details). The control group showed a normal inversion effect for matching face parts in accuracy ($t(10) = 1.746$, $p < .05$) and in response time ($t(10) = 4.754$, $p < .001$). However, they showed a paradoxical inversion effect for matching house-parts in accuracy

($t(10) = 1.743$, $p < .05$) and response time ($t(10) = 2.667$, $p < .01$). AM showed lower accuracies for matching both upright ($Z = -11.81$) and inverted ($Z = -5.36$) face-parts. Her latencies for matching upright face-parts ($Z = 2.51$) and house-parts ($Z = 2.06$) were higher than normal. She displayed a paradoxical inversion effect in the accuracy data for matching face-parts and house parts, and in the response times for matching house-parts. Her response times for matching face-parts showed a normal inversion pattern. HV had a reduced accuracy for matching upright face-parts ($Z = -2.00$). He also had highly prolonged response times for upright faces ($Z = 13.46$) and to a lesser extend for inverted faces ($Z = 8.27$). Latencies for upright houses ($Z = 2.28$) and inverted houses ($Z = 3.31$) were also prolonged, but less than for faces. HV showed paradoxical inversion effects in both the accuracy and response times for face-part and house-part matching. LW's accuracy for matching upright ($Z = -2.76$) and inverted ($Z = -3.16$) faces was impaired. His responses for matching upright face-parts ($Z = 8.87$), inverted face-parts ($Z = 5.13$), upright house-parts ($Z = 4.12$) and inverted house-parts ($Z = 4.61$) were prolonged. LW's accuracy data showed a normal inversion pattern for matching face-parts and a paradoxical inversion pattern for matching house parts. He displayed a paradoxical inversion effect in his response times for matching face-parts and house-parts.

fMRI analysis

All participants performed flawlessly on the oddball detection task.

Preprocessing

Imaging data were analyzed using Brainvoyager QX (Brain Innovation, Maastricht, the Netherlands). The first five volumes of each functional run were discarded to allow for T1 equilibration. Pre-processing of the functional data included 3D-motion correction, slice scan time correction, temporal data smoothing (high pass filter 3 cycles in time course) and spatial smoothing with an isotropic 6-mm full-width-half-maximum (FWHM) Gaussian kernel. Images were spatially normalized to Talairach space (Talairach & Tournoux, 1988) and resampled to a voxel size of 1 x 1 x 1 mm. Statistical analysis was based on the general linear model (GLM), with each condition defined as a separate predictor. Null-events were modeled explicitly.

ROI definition

We used a “split-half” method for defining regions of interest (ROI), in order to be sure that the observed effects are not due to a selection bias (Baker, Hutchison, & Kanwisher, 2007). The even trials were used to define the ROIs and the odd trials were used for the within ROI analysis. To localize face-sensitive activation in FG, i.e. FFA, we contrasted the even trials of all face conditions (fearful, happy and neutral) with houses (all trials) and identified significant voxels in each subject within a restricted region of the FG (Talairach y-coordinate between -25 and -65). The voxel set comprising this activation determined the ROI, in this case the FFA. The same procedure was followed in a restricted region of the IOG (Talairach y-coordinate < -70). To identify body sensitive areas, we compared the even trials of all bodies (fearful, happy and instrumental) with houses and mapped the selective activation in a restricted region of FG to determine the FBA (Talairach y-coordinate between -25 and -65) and the region around the junction of the middle temporal and middle occipital gyrus to determine the EBA (Talairach x-coordinate between 25 and 60; y-coordinate between -55 and -75; z-coordinate between -15 and 15). We used a liberal threshold ($p < .05$, uncorrected). Since previous studies reported that cortical face and body selective regions are often weaker or even absent in the left hemisphere (Kanwisher et al., 1997; Peelen & Downing, 2005), we restricted the analysis to the right hemisphere.

Smoothed activation maps are projected on the inflated right hemisphere of one subject. For every ROI, the activation maps of the control subjects are collapsed and the result is displayed by the black contours. This procedure allows visualization of the spatial extent of the activation across different subjects. Activation of the individual DPs is plotted in color (see Figures 2 to 5). The Talairach coordinates of the activation maps are shown in Table 3.

Table 3: Number of voxels (N) and Talairach coordinates (range) of ROIs. The coordinate range represents the outer voxels of the collapsed cluster from all controls. Abbreviations: n=neutral; f=fearful; h=happy; i=instrumental; F=face; B=body; H=house; FFA=Fusiform Face Area; IOG=Inferior Occipital Gyrus; FBA=Fusiform Body Area; EBA=Extrastriate Body Area; AMG=Amygdala

Contrast	Area	Mean Controls (range)						AM				HV				LW			
		N	x	y	z	N	x	y	z	N	x	y	z	N	x	y	z		
(fF+hF+nF)>H																			
	FFA	427 (9:2208)	40 (31:48)	-42 (-59:-27)	-15 (-33:-5)	24	39	-60	-20	448	36	-64	-9	140	41	-34	-16		
	IOG	742 (15:2667)	32 (15:55)	-81 (-95:-70)	-7 (-21:9)	9	24	-82	-14	534	25	-74	-13	156	47	-71	-10		
(fB+hB+iB)>H																			
	FBA	1220 (303:2282)	44 (27:58)	-40 (-60:-25)	-15 (-23:-19)	406	34	-49	-9	502	41	-40	-18	34	41	-54	-11		
	EBA	416 (42:1201)	45 (33:59)	-63 (-74:-52)	4 (-11:15)	737	43	-70	-6	802	36	-66	-10	2404	43	-68	8		
fF>nF																			
	Right AMG	146(146:146)	12(8:16)	-8(-4:-12)	-12(-14:-10)					14	24	0	-8	20	23	0	-7		
	Left AMG					15	-20	-8	-3	80	-24	1	-5						
hF>nF																			
	Right AMG	11(0:11)	9(8:9)	-1(-2:0)	-11(-12:-10)									23	23	0	-8		
	Left AMG	265(0:265)	-13(-18:-11)	-12(-14:-9)	-14(-18:-19)					35	-24	-1	-6						
fB>nB																			
	Right AMG	258(8:779)	19 (12:26)	-7 (-9:0)	-12 (-17:-7)					142	13	-6	-6						
	Left AMG	351 (0:710)	-20 (-28:-11)	-8 (-9:0)	-10 (-19:18)	130	-23	-1	-18	120	-17	-10	-11						
hB>nB																			
	Right AMG	152 (0:330)	19 (13:25)	-5 (-14:0)	-16 (-21:-7)														
	Left AMG	659 (0:819)	-17 (-24:-11)	-4 (-11:0)	-13 (-21:-7)					102	-19	-11	-6						

Effects of emotional content

The analyses were performed on the beta-values of the odd trials of the conditions. To investigate differences between the DP group and the control group, we used independent samples t-tests, corrected for unequal variances (in degrees of freedom).

FFA

Figure 2 shows the smoothed face-specific activation (left) and the beta-values of all conditions (right) in FG. The controls show the expected age-dependend higher activation for neutral than for fearful expressions (Guyer et al., 2008). We calculated the difference between fearful faces and neutral faces and this difference was significantly larger in the control group ($t(4.946) = -2.583, p < .05$). The difference between happy faces and neutral faces was marginally significantly different between groups ($t(4.906) = -2.051, p < .097$). Since previous studies showed a lower activation for faces in DPs compared to controls (Bentin et al., 2007; Hadjikhani & de Gelder, 2002), we used one-tailed post-hoc t-tests to compare the activation levels of the three face conditions between both groups. This revealed a marginally significant difference for the neutral faces ($t(4.980) = 1.929, p < .051$).

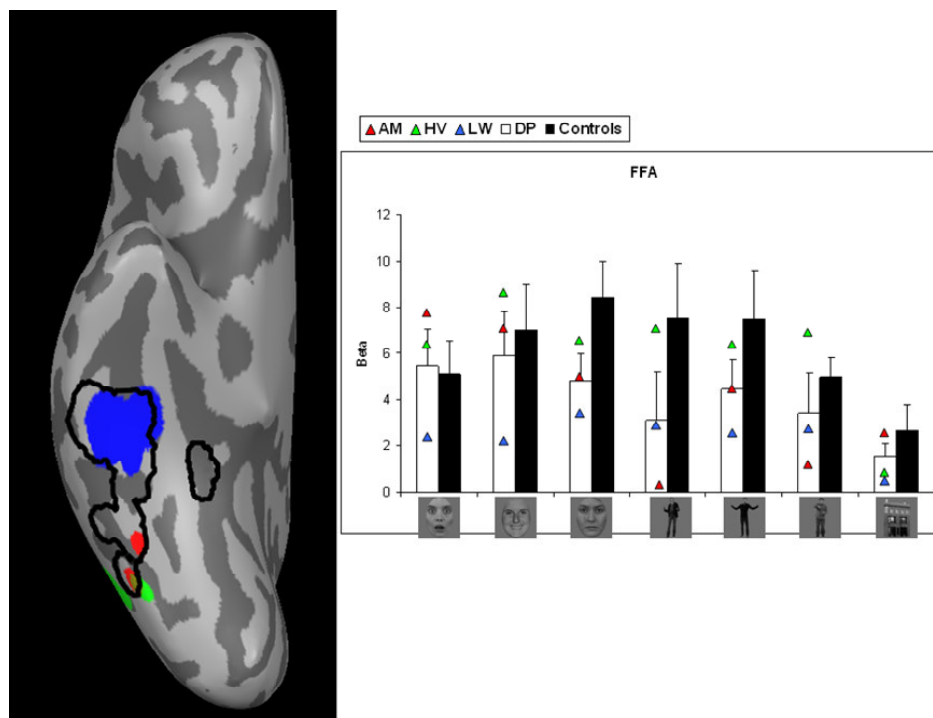


Figure 2. Face-specific activation in right FG when comparing faces (fearful/happy/neutral) with houses. Left: Areas are shown on an inflated right hemisphere. Activation maps of the control subjects are collapsed and displayed by the black contours. Activation of the individual DPs is plotted in color. Right: beta-values by condition, group and DP. Error bars represent one standard error of the mean (SEM). Conditions represent from left to right: fearful faces, happy faces, neutral faces, fearful bodies, happy bodies, neutral bodies and houses. White columns display the average value of the three patients. Black columns show the average value of the controls. Triangles represent the individual values of the DPs.

IOG

Figure 3 shows the smoothed face-specific activation (left) in IOG and the beta-values of all conditions (right). A t-test on the difference between fearful faces and neutral faces showed no significant difference between both groups ($t(4.510) = .0233$, $p < .826$). The difference between happy faces and neutral faces was also not significantly different between the DPs and controls ($t(4.989) = -1.235$, $p < .272$).

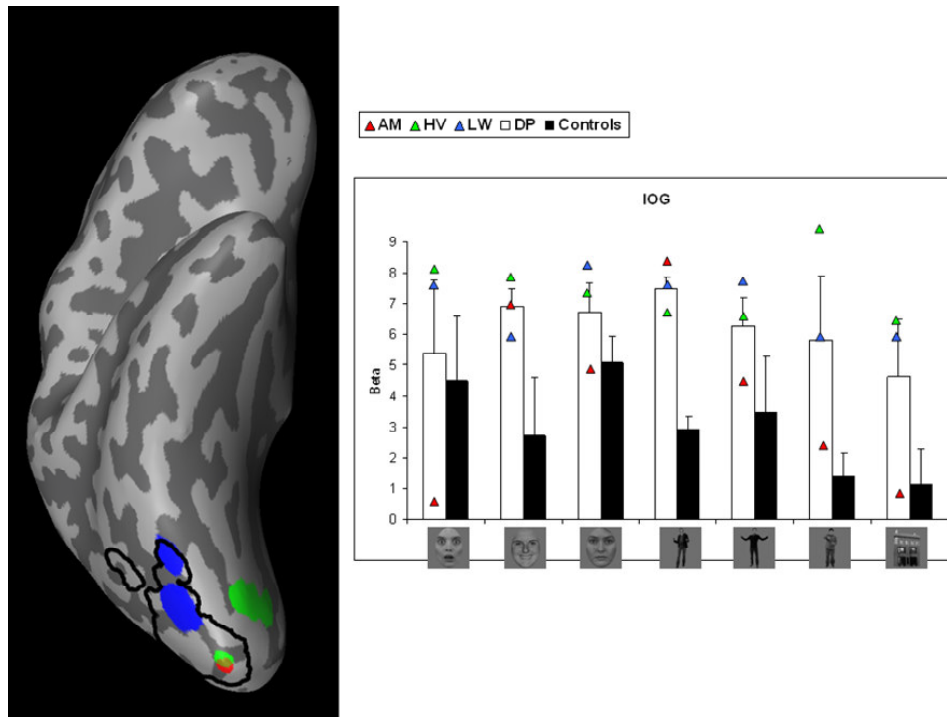


Figure 3. Face-specific activation in right IOG when comparing faces (fearful/happy/neutral) with houses. Left: Areas are shown on an inflated right hemisphere. Activation maps of the control subjects are collapsed and displayed by the black contours. Activation of the individual DPs is plotted in color. Right: beta-values by condition, group and DP. Error bars represent one SEM. Conditions represent from left to right: fearful faces, happy faces, neutral faces, fearful bodies, happy bodies, neutral bodies and houses. White columns display the average value of the three patients. Black columns show the average value of the controls. Triangles represent the individual values of the DPs.

FBA

Figure 4 shows the smoothed body-specific activation (left) and the beta-values of all conditions (right) in FBA. The difference between either fearful bodies ($t(4.475) = -.088$, $p < .934$) or happy bodies ($t(4.567) = .321$, $p < .762$) and instrumental bodies was not significantly different between both groups.

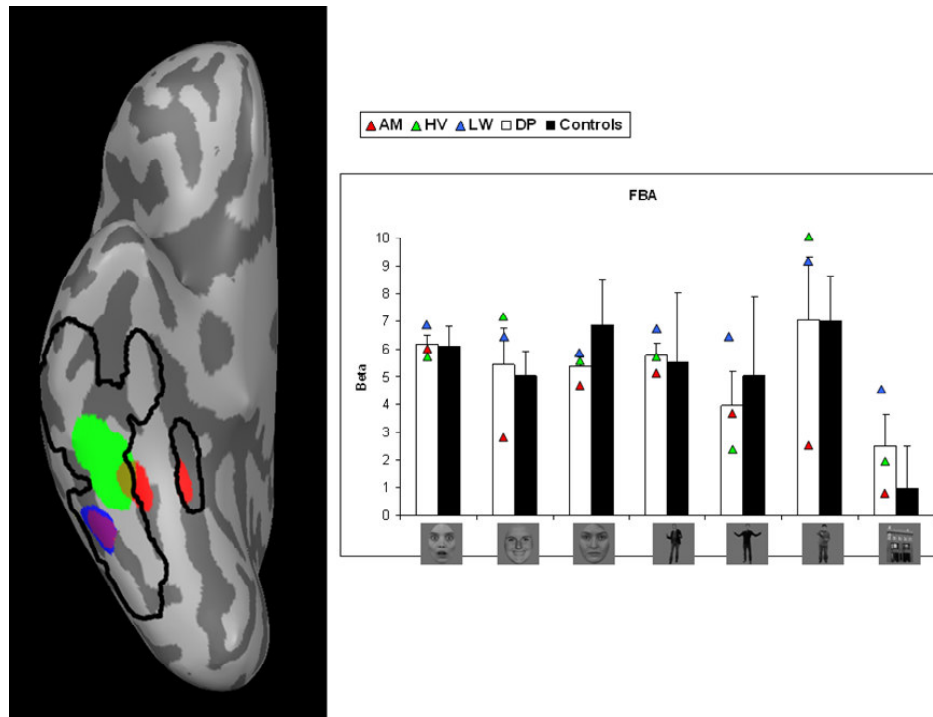


Figure 4. Body-specific activation in right FG when comparing bodies (fearful/happy/instrumental) with houses. Left: Areas are shown on an inflated right hemisphere. Activation maps of the control subjects are collapsed and displayed by the black contours. Activation of the individual DPs is plotted in color. The purple indicates overlap between red (AM) and blue (LW). Right: beta-values by condition, group and DP. Error bars represent one SEM. Conditions represent from left to right: fearful faces, happy faces, neutral faces, fearful bodies, happy bodies, neutral bodies and houses. White columns display the average value of the three patients. Black columns show the average value of the controls. Triangles represent the individual values of the DPs.

EBA

Figure 5 shows the smoothed body-specific activation (left) and the beta-values of all conditions (right) in EBA. The difference between fearful bodies and instrumental bodies was not different between groups ($t(3.786) = 1.153$, $p < .317$). A t- test on the difference between happy and instrumental bodies revealed no significant between-group difference ($t(3.722) = .339$, $p < .573$).

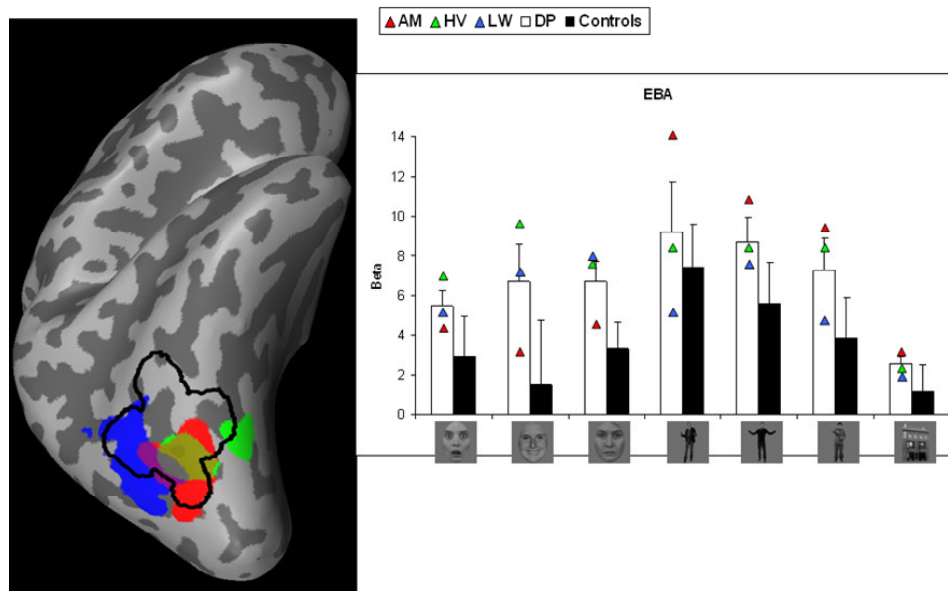


Figure 5. Body-specific activation in right EBA when comparing bodies (fearful/happy/instrumental) with houses. Left: Areas are shown on an inflated right hemisphere. Activation maps of the control subjects are collapsed and displayed by the black contours. Activation of the individual DPs is plotted in color. The purple indicates overlap between red (AM) and blue (LW). Right: beta-values by condition, group and DP. Error bars represent one SEM. Conditions represent from left to right: fearful faces, happy faces, neutral faces, fearful bodies, happy bodies, neutral bodies and houses. White columns display the average value of the three patients. Black columns show the average value of the controls. Triangles represent the individual values of the DPs.

Effects of categorical selectivity

To investigate the selectivity of processing faces and bodies in the brain, we calculated the difference between the mean of the three face conditions and the mean of the three body conditions in FFA and IOG. A comparison using t-tests showed that this difference was smaller in the control group in IOG, but it did not reach statistical significance ($t(3.961) = 2.122$, $p < .102$). We also calculated the difference between the mean of all body conditions and the mean of all face conditions in FBA and EBA. Independent sample t-tests showed no significant between-group differences.

Processing of neutral faces

Since the main body of research on DP concerns neutral faces, we compared the activation level of neutral faces between both groups in all four ROIs, using t-tests. In addition to the above mentioned difference in FFA, this revealed a marginally significantly higher activation for neutral faces in EBA in the DP group ($t(4.955) = 2.044$, $p < .097$).

Effects of emotion in amygdala

Finally, we performed a post-hoc analysis, in which we defined the amygdala in each subject, based on the individual anatomy. This ROI consisted in each hemisphere of a cube of $13 \times 13 \times 13$ voxels around the center of the amygdala and we performed a second GLM in this area. The results are shown in Table 3. Contrasting fearful faces with neutral faces revealed significant activation in all three patients (left amygdala in AM; bilateral amygdala in HV and right amygdala in LW). Comparing happy with neutral faces showed activation in two patients (left amygdala in HV and right amygdala in LW). Fearful compared with neutral bodies differentially activated the amygdala in two patients (left amygdala in AM and bilateral amygdala HV). Happy bodies triggered significantly more amygdala activity in one DP (left amygdala in HV) compared to neutral bodies.

Discussion

The first major finding is that compared to the control group, the DP group displays a similar activation level for the emotional faces, but a lower activation in FFA for neutral faces. A lower activation level in DP for neutral face perception in FG is consistent with earlier reports (Bentin et al., 2007; Hadjikhani & de Gelder, 2002). The present results are compatible with the theoretical perspective on face recognition difficulties argued for previously (Damasio et al., 1982; Damasio et al., 1990) suggesting a higher threshold for neutral face recognition performance in prosopagnosics. This relative difficulty with neutral faces is based on the notion that faces are more difficult stimuli than many other categories they are routinely compared with.

Emotional stimuli trigger a higher level of arousal (e.g. Lang, Greenwald, Bradley, & Hamm, 1993; Mehrabian & Russell, 1974) and emotion in a face constitutes an additional feature that carries important communicative information and is therefore more salient. This saliency hypothesis is supported by a number of behavioral studies, with different visual tasks, that have demonstrated that adding emotional information to a face results in a greater tendency to capture attention (Eastwood, Smilek, & Merikle, 2003; Fox, Russo, Bowles, & Dutton, 2001; Pourtois, Grandjean, Sander, & Vuilleumier, 2004; Vuilleumier & Schwartz, 2001). Note though that the emotion effects we observe are not specific for emotions with a negative valence since we obtain similar effects for both fearful and happy (although less pronounced) expressions.

However, normal FFA activation for facial expressions in the presence of lower than normal activation for neutral faces suggests that the activation boost is triggered more in the emotion processing than in the impaired face processing system in ventro-temporal cortex. Studies on perception of emotional faces in normals have hypothesized the existence of a feedback mechanism between FG and amygdala (Rossion et al., 2003; Rotshtein et al., 2001; Surguladze et al., 2003; Vuilleumier et al., 2001; Vuilleumier, Richardson, Armony, Driver, & Dolan, 2004). The possibility that such feedback connections from the amygdala may be active in prosopagnosia and boost face processing was already suggested in an earlier study of emotional faces in prosopagnosia (de Gelder et al., 2003). Two acquired prosopagnosics were presented with both a neutral and emotional part-to-whole face matching task. The patients had lesions in FG and/or IOG, but the results showed normal activation in other face-sensitive areas (amygdala, superior temporal sulcus), for the contrast between emotional and neutral faces. The patients were also more accurate and faster when they performed the task with emotional faces compared to neutral ones. Moreover, the patients showed a normal inversion effect for matching emotional but not for neutral faces.

Lower neural activity in the DPs for neutral faces, but not for emotional faces is compatible with a dual route model of face perception as argued first in de Gelder and Rouw (B. de Gelder & R. Rouw, 2000a) and adapted in de Gelder et al. (de Gelder et al., 2003), involving subcortical structures along a pathway that is able to process facial expressions (the pulvinar-superior colliculus-amygdala route) (Morris, Ohman, & Dolan, 1999) which in turn may boost face representations in the cortical route in temporal cortex even when face representations in temporal cortex are weak as shown by the lower activation for neutral faces in the DP group (de Gelder et al., 2003). The pattern observed here is in line with this and may also explain

why emotional content facilitates the cortical processing of faces in prosopagnosia. Consistent with this, we observed a higher activity level of the amygdala for emotional faces compared to neutral ones. A related and more extreme phenomenon is observed in hemianopic patients, who are unable to consciously report the presentation of a face in the blind visual field and do not show FG activation when presented facial expressions in the blind field but who perform well above chance in tasks where they have to guess the facial expression (de Gelder, Vroomen, Pourtois, & Weiskrantz, 1999).

Our second main finding concerns the categorical specificity of face vs. body representation in DPs. We compared the activation of body conditions in the face selective regions and of the face conditions in the body selective regions between both groups. On the one hand, our findings indicate that perceiving neutral faces results in a higher activation of EBA in the DP group, compared to the control group. Combined with the lower activation for neutral faces in FFA, this increased activation in EBA might indicate an anomalous cerebral processing route in DP. It may be the case that (neutral) faces are processed in the areas more dominantly dedicated to body perception. On the other hand, we find a higher activation for perceiving bodies in IOG. These combined findings indicate that the neural correlates of perceiving faces and bodies, as manifested in IOG and EBA show a lower degree of specificity in DP.

For body triggered activity we find no difference in neutral vs. emotional expressions between both groups, either in FBA or EBA. This indicates that the anomalous neuro-functional substrate in our DP group for neutral faces does not extend to the processing of bodies and bodily expressions. This is in line with recent behavioral data showing no impairment in recognizing neutral body postures in one DP patient (Duchaine et al., 2006). One of the DPs (HV) in the present study participated in a previous ERP study on perception of neutral faces and neutral bodies (Righart & de Gelder, 2007) and the results of both studies are partly converging. Righart & de Gelder (2007) measured the electrical brain correlates of the inversion effect as an index of configural processes (the ability to perceive stimuli as one configuration as opposed to an assemblage of features (Young, Hellawell, & Hay, 1987)). HV differed significantly from the control group in face processing on two accounts. He displayed a paradoxical ERP inversion effect (the reverse pattern from the controls) around 100 ms after stimulus presentation (P1 amplitude) and no inversion effect around 170 ms after stimulus presentation (N170 latency). But his results for bodies did not differ from the controls.

An important and relevant difference between face and body perception concerns the coding of identity. A face contains all necessary information about the identity of a person and we are used and trained to recognize identity by the face. A person can be readily identified on the basis of his face, but identification based on the body alone is far less evident. The different pattern in FG for faces and bodies may therefore reflect the possibility that FG is more involved in processing person identity (Grill-Spector et al., 2004) which is typically more based on the face than on the body.

Notwithstanding the well documented involvement of FG in face perception, its precise role of FG in prosopagnosia is still a matter of debate. We do not clearly understand at present how factors like maturation of different cortical areas, like the FG, are important for normal face recognition. Reduced volume of the right temporal lobe has previously been reported in a DP patient (Bentin, Deouell, & Soroker, 1999). A structural imaging study in six DP subjects investigated volumetric and morphometric properties in occipito-temporal cortex and showed a decreased volume of the FG that correlated with face recognition deficits (Behrmann, Avidan, Gao, & Black, 2007). At the neuro-functional level, recent data collected from normals show a correlation between the volumetric size of the right FFA and recognition memory for neutral faces (Golarai et al., 2007). This study also investigated the development of category specific brain areas and the results suggest that the relative size of the FFA increases during development. Moreover, the development of the FFA takes longer compared to that of object selective areas (lateral occipital complex) or face sensitive areas in the superior temporal sulcus (see Grill-Spector, Golarai, & Gabrieli, 2008 for review and discussion). These findings support the notion that DP may be associated with abnormal development of FG which may be either a consequence or a cause of anomalous face skills. Lesions in acquired prosopagnosia (AP) patients often include the FG (Barton, Press, Keenan, & O'Connor, 2002; e.g. de Gelder et al., 2003), although other cases have also been reported with lesions more posterior than the face sensitive part of the FG (e.g. Sorger, Goebel, Schiltz, & Rossion, 2007; Steeves et al., 2006). Besides the heterogeneity across lesion localization in AP, considerable heterogeneity consists in behavioral symptoms in DP (Le Grand et al., 2006). Since successful face-processing is likely to involve a variety of hierarchical and parallel processes, impairments in different processes will result in different types of

behavioral and neuro-anatomical correlates. The results from the present study clearly demonstrate the importance of emotional information in face processing and urge (future imaging) studies to take the modulatory effect of emotion into account, in order to further untangle the complex nature of DP.

References

- Aguirre, G. K., Zarahn, E., & D'Esposito, M. (1998). An area within human ventral cortex sensitive to "building" stimuli: evidence and implications. *Neuron*, 21(2), 373-383.
- Avidan, G., Hasson, U., Malach, R., & Behrmann, M. (2005). Detailed exploration of face-related processing in congenital prosopagnosia: 2. Functional neuroimaging findings. *J Cogn Neurosci*, 17(7), 1150-1167.
- Baker, C. I., Hutchison, T. L., & Kanwisher, N. (2007). Does the fusiform face area contain subregions highly selective for nonfaces? *Nat Neurosci*, 10(1), 3-4.
- Barton, J. J., Press, D. Z., Keenan, J. P., & O'Connor, M. (2002). Lesions of the fusiform face area impair perception of facial configuration in prosopagnosia. *Neurology*, 58(1), 71-78.
- Behrmann, M., Avidan, G., Gao, F., & Black, S. (2007). Structural imaging reveals anatomical alterations in inferotemporal cortex in congenital prosopagnosia. *Cereb Cortex*, 17(10), 2354-2363.
- Bentin, S., Degutis, J. M., D'Esposito, M., & Robertson, L. C. (2007). Too Many Trees to See the Forest: Performance, Event-related Potential, and Functional Magnetic Resonance Imaging Manifestations of Integrative Congenital Prosopagnosia. *J Cogn Neurosci*, 19(1), 132-146.
- Bentin, S., Deouell, L. Y., & Soroker, N. (1999). Selective visual streaming in face recognition: evidence from developmental prosopagnosia. *Neuroreport*, 10(4), 823-827.
- Benton, A. L., Sivan, A. B., Hamsher, K., Varney, N. R., & Spreen, O. (1983). Contribution to neuropsychological assessment. NY: Oxford University Press.
- Bodamer, J. (1947). Die prosop-Agnosie. *Archiv fur Psychiatrie und Nervenkrankheiten*, 179, 6-53.
- Breiter, H. C., Etcoff, N. L., Whalen, P. J., Kennedy, W. A., Rauch, S. L., Buckner, R. L., et al. (1996). Response and Habituation Of the Human Amygdala During Visual Processing Of Facial Expression. *Neuron*, 17(5), 875-887.
- Damasio, A. R., Damasio, H., & Van Hoesen, G. W. (1982). Prosopagnosia: anatomic basis and behavioral mechanisms. *Neurology*, 32(4), 331-341.
- Damasio, A. R., Tranel, D., & Damasio, H. (1990). Face agnosia and the neural substrates of memory. *Annu Rev Neurosci*, 13, 89-109.

- de Gelder, B. (2006). Towards the neurobiology of emotional body language. *Nature Reviews Neuroscience*, 7(3), 242-249.
- de Gelder, B., Bachoud-Levi, A. C., & Degos, J. D. (1998). Inversion superiority in visual agnosia may be common to a variety of orientation polarised objects besides faces. *Vision Res*, 38(18), 2855-2861.
- de Gelder, B., Frissen, I., Barton, J., & Hadjikhani, N. (2003). A modulatory role for facial expressions in prosopagnosia. *Proc Natl Acad Sci U S A*.
- de Gelder, B., & Rouw, P. (2000). Structural encoding precludes recognition of parts in prosopagnosia. *Cognitive Neuropsychology*, 17, 89-102.
- de Gelder, B., & Rouw, R. (2000a). Configural face processes in acquired and developmental prosopagnosia: evidence for two separate face systems? *Neuroreport*, 11(14), 3145-3150.
- de Gelder, B., & Rouw, R. (2000b). Paradoxical configuration effects for faces and objects in prosopagnosia. *Neuropsychologia*, 38(9), 1271-1279.
- de Gelder, B., Vroomen, J., Pourtois, G., & Weiskrantz, L. (1999). Non-conscious recognition of affect in the absence of striate cortex. *Neuroreport*, 10(18), 3759-3763.
- Degutis, J. M., Bentin, S., Robertson, L. C., & D'Esposito, M. (2007). Functional Plasticity in Ventral Temporal Cortex following Cognitive Rehabilitation of a Congenital Prosopagnosic. *J Cogn Neurosci*, 19(11), 1790-1802.
- Diamond, R., & Carey, S. (1986). Why faces are and are not special: an effect of expertise. *J Exp Psychol Gen*, 115(2), 107-117.
- Dolan, R. J., Fletcher, P. C., Morris, J. S., Kapur, N., Deakin, J. F., & Frith, C. D. (1996). Neural activation during covert processing of positive emotional facial expressions. *Neuroimage*, 4(3 Pt 1), 194-200.
- Dolan, R. J., Morris, J. S., & de Gelder, B. (2001). Crossmodal binding of fear in voice and face. *Proc Natl Acad Sci U S A*, 98(17), 10006-10010.
- Downing, P. E., Jiang, Y., Shuman, M., & Kanwisher, N. (2001). A cortical area selective for visual processing of the human body. *Science*, 293(5539), 2470-2473.
- Duchaine, B. C., Parker, H., & Nakayama, K. (2003). Normal recognition of emotion in a prosopagnosic. *Perception*, 32(7), 827-838.

- Duchaine, B. C., Yovel, G., Butterworth, E. J., & Nakayama, K. (2006). Prosopagnosia as an impairment to face-specific mechanisms: Elimination of the alternative hypotheses in a developmental case. *Cognitive Neuropsychology*, 23(5), 714-747.
- Eastwood, J. D., Smilek, D., & Merikle, P. M. (2003). Negative facial expression captures attention and disrupts performance. *Percept Psychophys*, 65(3), 352-358.
- Epstein, R., & Kanwisher, N. (1998). A cortical representation of the local visual environment. *Nature*, 392(6676), 598-601.
- Farah, M. (1990). *Visual agnosia: Disorders of visual recognition and what they tell us about normal vision*. Cambridge: MIT Press.
- Fodor, J. (1983). *The Modularity of Mind*. Cambridge, MA: MIT Press.
- Fox, E., Russo, R., Bowles, R., & Dutton, K. (2001). Do threatening stimuli draw or hold visual attention in subclinical anxiety? *J Exp Psychol Gen*, 130(4), 681-700.
- Friston, K. J., Zarahn, E., Josephs, O., Henson, R. N., & Dale, A. M. (1999). Stochastic designs in event-related fMRI. *Neuroimage*, 10(5), 607-619.
- Gauthier, I., Behrmann, M., & Tarr, M. J. (1999). Can face recognition really be dissociated from object recognition? *J Cogn Neurosci*, 11, 349-370.
- Gauthier, I., Skudlarski, P., Gore, J. C., & Anderson, A. W. (2000). Expertise for cars and birds recruits brain areas involved in face recognition. *Nature Neuroscience*, 3(2), 191-197.
- Golarai, G., Ghahremani, D. G., Whitfield-Gabrieli, S., Reiss, A., Eberhardt, J. L., Gabrieli, J. D., et al. (2007). Differential development of high-level visual cortex correlates with category-specific recognition memory. *Nat Neurosci*, 10(4), 512-522.
- Grill-Spector, K., Golarai, G., & Gabrieli, J. (2008). Developmental neuroimaging of the human ventral visual cortex. *Trends Cogn Sci*, 12(4), 152-162.
- Grill-Spector, K., Knouf, N., & Kanwisher, N. (2004). The fusiform face area subserves face perception, not generic within-category identification. *Nat Neurosci*, 7(5), 555-562.
- Grueter, M., Grueter, T., Bell, V., Horst, J., Laskowski, W., Sperling, K., et al. (2007). Hereditary prosopagnosia: the first case series. *Cortex*, 43(6), 734-749.
- Guyer, A. E., Monk, C. S., McClure-Tone, E. B., Nelson, E. E., Roberson-Nay, R., Adler, A. D., et al. (2008). A Developmental Examination of Amygdala Response to Facial Expressions. *J Cogn Neurosci*.

- Hadjikhani, N., & de Gelder, B. (2002). Neural basis of prosopagnosia: an fMRI study. *Hum Brain Mapp*, 16(3), 176-182.
- Hadjikhani, N., & de Gelder, B. (2003). Seeing fearful body expressions activates the fusiform cortex and amygdala. *Curr Biol*, 13(24), 2201-2205.
- Hasson, U., Avidan, G., Deouell, L. Y., Bentin, S., & Malach, R. (2003). Face-selective activation in a congenital prosopagnosic subject. *Journal of Cognitive Neuroscience*, 15(3), 419-431.
- Haxby, J. V., Hoffman, E. A., & Gobbini, M. I. (2000). The distributed human neural system for face perception. *Trends Cogn Sci*, 4(6), 223-233.
- Hoffman, E. A., & Haxby, J. V. (2000). Distinct representations of eye gaze and identity in the distributed human neural system for face perception. *Nat Neurosci*, 3(1), 80-84.
- Johnson-Frey, S. H., Maloof, F. R., Newman-Norlund, R., Farrer, C., Inati, S., & Grafton, S. T. (2003). Actions or hand-object interactions? Human inferior frontal cortex and action observation. *Neuron*, 39(6), 1053-1058.
- Jones, R. D., & Tranel, D. (2001). Severe developmental prosopagnosia in a child with superior intellect. *J Clin Exp Neuropsychol*, 23(3), 265-273.
- Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: a module in human extrastriate cortex specialized for face perception. *J Neurosci*, 17(11), 4302-4311.
- Kennerknecht, I., Grueter, T., Welling, B., Wentzek, S., Horst, J., Edwards, S., et al. (2006). First report of prevalence of non-syndromic hereditary prosopagnosia (HPA). *Am J Med Genet A*, 140(15), 1617-1622.
- Lang, P. J., Greenwald, M. K., Bradley, M. M., & Hamm, A. O. (1993). Looking at pictures: Affective, facial, visceral and behavioral reactions. *Psychophysiology*, 30, 261-273.
- Le Grand, R., Cooper, P. A., Mondloch, C. J., Lewis, T. L., Sagiv, N., de Gelder, B., et al. (2006). What aspects of face processing are impaired in developmental prosopagnosia? *Brain and Cognition*, 61(2), 139-158.
- Levy, I., Hasson, U., Avidan, G., Hendler, T., & Malach, R. (2001). Center-periphery organization of human object areas. *Nat Neurosci*, 4(5), 533-539.
- Lundqvist, D., Flykt, A., & Öhman, A. (1998). *The Karolinska Directed Emotional Faces - KDEF*. Stockholm: Karolinska Institutet.

- Mehrabian, A., & Russell, J. A. (1974). *An approach to environmental psychology*. Cambridge, MA: MIT Press.
- Morris, J. S., Ohman, A., & Dolan, R. J. (1999). A subcortical pathway to the right amygdala mediating "unseen" fear. *Proc Natl Acad Sci U S A*, 96(4), 1680-1685.
- Nunn, J. A., Postma, P., & Pearson, R. (2001). Developmental prosopagnosia: should it be taken at face value? *Neurocase*, 7(1), 15-27.
- Peelen, M. V., & Downing, P. E. (2005). Selectivity for the human body in the fusiform gyrus. *J Neurophysiol*, 93(1), 603-608.
- Peelen, M. V., & Downing, P. E. (2007). The neural basis of visual body perception. *Nat Rev Neurosci*, 8(8), 636-648.
- Pourtois, G., Grandjean, D., Sander, D., & Vuilleumier, P. (2004). Electrophysiological correlates of rapid spatial orienting towards fearful faces. *Cereb Cortex*, 14(6), 619-633.
- Puce, A., Allison, T., Asgari, M., Gore, J. C., & McCarthy, G. (1996). Differential sensitivity of human visual cortex to faces, letterstrings, and textures: a functional magnetic resonance imaging study. *J Neurosci*, 16(16), 5205-5215.
- Reed, C. L., Stone, V. E., Bozova, S., & Tanaka, J. (2003). The body-inversion effect. *Psychol Sci*, 14(4), 302-308.
- Riddoch, M. J., & Humphreys, G. W. (1993). *Birmingham Object Recognition Battery*. Hove: Psychology Press.
- Righart, R., & de Gelder, B. (2007). Impaired face and body perception in developmental prosopagnosia. *Proc Natl Acad Sci U S A*.
- Rossion, B., Caldara, R., Seghier, M., Schuller, A. M., Lazeyras, F., & Mayer, E. (2003). A network of occipito-temporal face-sensitive areas besides the right middle fusiform gyrus is necessary for normal face processing. *Brain*, 126, 2381-2395.
- Rotshtein, P., Malach, R., Hadar, U., Graif, M., & Hendler, T. (2001). Feeling or features: different sensitivity to emotion in high-order visual cortex and amygdala. *Neuron*, 32(4), 747-757.
- Sorger, B., Goebel, R., Schiltz, C., & Rossion, B. (2007). Understanding the functional neuroanatomy of acquired prosopagnosia. *Neuroimage*, 35(2), 836-852.

- Steeves, J. K., Culham, J. C., Duchaine, B. C., Pratesi, C. C., Valyear, K. F., Schindler, I., et al. (2006). The fusiform face area is not sufficient for face recognition: evidence from a patient with dense prosopagnosia and no occipital face area. *Neuropsychologia*, 44(4), 594-609.
- Stekelenburg, J. J., & de Gelder, B. (2004). The neural correlates of perceiving human bodies: an ERP study on the body-inversion effect. *Neuroreport*, 15(5), 777-780.
- Surguladze, S. A., Brammer, M. J., Young, A. W., Andrew, C., Travis, M. J., Williams, S. C., et al. (2003). A preferential increase in the extrastriate response to signals of danger. *Neuroimage*, 19(4), 1317-1328.
- Talairach, J., & Tournoux, P. (1988). *Co-Planar Stereotaxic Atlas of the Human Brain*. New York: Thieme Medical Publishers.
- van de Riet, W. A. C., Grèzes, J., & de Gelder, B. (in press). Specific and common brain regions involved in the perception of faces and bodies and the representation of their emotional expressions. *Social Neuroscience*.
- Vuilleumier, P., Armony, J. L., Driver, J., & Dolan, R. J. (2001). Effects of attention and emotion on face processing in the human brain: an event-related fMRI study. *Neuron*, 30(3), 829-841.
- Vuilleumier, P., Richardson, M. P., Armony, J. L., Driver, J., & Dolan, R. J. (2004). Distant influences of amygdala lesion on visual cortical activation during emotional face processing. *Nat Neurosci*, 7(11), 1271-1278.
- Vuilleumier, P., & Schwartz, S. (2001). Emotional facial expressions capture attention. *Neurology*, 56, 153-158.
- Warrington, E. K. (1984). *Recognition Memory Test*. Nelson, Windsor: NFER.
- Williams, M. A., Berberovic, N., & Mattingley, J. B. (2007). Abnormal fMRI adaptation to unfamiliar faces in a case of developmental prosopamnesia. *Curr Biol*, 17(14), 1259-1264.
- Yin, R. K. (1969). Looking at upside-down faces. *Journal of Experimental Psychology*, 81, 141-145.
- Young, A. W., Hellawell, D., & Hay, D. C. (1987). Configurational information in face perception. *Perception*, 16, 747-759.

Chapter 10: Huntington's disease impairs recognition of angry and instrumental body language¹⁰

Introduction

Huntington's disease (HD) attacks primarily basal ganglia structures (mostly caudate nucleus and putamen) leading to severe motor deficits (Vonsattel *et al.*, 1985). At the same time HD is accompanied by deficits in recognizing emotional expressions, particularly facial expressions of disgust (Sprengelmeyer et al., 1996; Wang, Hoosain, Yang, Meng, & Wang, 2003).

However, in natural circumstances facial expressions are rarely seen in isolation, but occur in the context of expressions by the whole body. Our ability to perceive these emotional body expressions and their representation in the brain is now becoming an important research topic (de Gelder, 2006), and there are strong arguments for exploring how HD patients recognize emotional body language. In previous studies with neurologically intact observers, we used fMRI to clarify how the brain recognizes happiness or fear expressed by a whole body (de Gelder, Snyder, Greve, Gerard, & Hadjikhani, 2004; Grezes, Pichon, & de Gelder, 2007). Our results indicate that observing fearful body expressions produces increased activity in brain areas associated with perception of emotional faces, but also in areas involved in representation of action and movement, including caudate nucleus and putamen. Caudate nucleus and putamen are known for their involvement in motor tasks but have also been associated with motivational and emotional task components (Bhatia & Marsden, 1994; Grillner, Hellgren, Menard, Saitoh, & Wikstrom, 2005; Kampe, Frith, Dolan, & Frith, 2001). Therefore, we conjectured that HD is also associated with a deficit in recognizing emotions expressed by the whole body. However, since our pilot data indicated that static bodily expressions of disgust are difficult to distinguish from fear (both postures consisting in moving backwards and putting hand palms forward), we did not include disgust.

¹⁰ This chapter is based on: de Gelder, B., Van den Stock, J., de Diego Balaguer, R., & Bachoud-Lévi, A.-C. (2008). Huntington's disease impairs recognition of angry and instrumental body language. *Neuropsychologia*, 46, 369-373.

Method

Participants. Nineteen HD patients (10 early HD at stage I and 9 at stage II, using the classification based on the Total Functional Capacity scale (Shoulson, 1981) and 19 control subjects participated. HD patients were recruited from an out-clinic follow-up program within the framework of interventional therapy approved by the ethical committee of the Henri Mondor Hospital. They had no previous neurological or psychiatric history and their HD diagnosis was genetically confirmed. All subjects gave informed consent. Patients were evaluated using the Mattis Dementia Rating Scale (MDRS) (Mattis, 1976), and the United Huntington's Disease Rating Scale of which the cognitive part contains the Stroop test, the verbal fluency task and Symbol Digit Test. All patients were administered by the same rater. Atrophy of the caudate was assessed in 11 patients with MRI by calculating an adjusted bicaudate ratio, which took cortical atrophy into account (the minimal distance between the caudate indentations of the frontal horns divided by the width of the brain along the same line multiplied by 100). We opted for this adjustment since there are now several studies showing rather widespread cortical pathology in HD gene carriers (Kassubek, Bernhard Landwehrmeyer et al., 2004; Kassubek, Gaus, & Landwehrmeyer, 2004; Thieben et al., 2002).

Control subjects were healthy volunteers with normal or corrected to normal vision and no previous neurological history. They were matched on age, $t(36) = 1.37, p = 0.180$, sex, $\chi^2 = 1.18, p = 0.669$, dexterity, $\chi^2 = 0.01, p = 0.958$ and years of education, $t(36) = 0.29, p = 0.775$ (see Table 1 for demographic and general assessment data).

Table 1 Demographic Data of HD Patients and Control Subjects. n. a. = not applicable *N=11. The norms are issued from % (Golden, 1978); § (Cardebat et al., 1990); μ (Wechsler, 1981); # (Starkstein et al., 1989).

Demographic data	HD	Controls
N	19	19
Sex	9F/ 10M	8F/ 11M
Age in years (SD)	52.0 (9.1)	48.2 (8.0)
Educational level in years (SD)	13.9 (4.6)	14.4 (5.1)
Evolution duration in years (SD)	5.9 (4.2)	n. a.
CAG-repeats (SD)	42.7 (2.6)	n. a.
Laterality	18R/ 1L	18R/ 1L
General assessment data	Normal published range	
Total Functional Capacity	10.0 (2.1)	13
UHDRS motor score	32.4 (19.0)	0
MDRS	129.8 (7.7)	≥136
Stroop Color/word	20.3 (10.0)	≥35%
Fluency “P” in 2 min	16.8 (7.1)	18 §
Symbol Digit Code	21.9 (8.6)	≥37 ^μ
Bicaudate ratio's*	20.7 (4.3)	< 10 #

Stimuli and procedure. Video recordings of eight semi-professional actors (half of them women, age 22-35 years) were used for stimulus construction. Actors performed fearful, angry and sad expressive gestures with their whole body. The actors also performed instrumental but emotionally neutral actions (pouring water into a glass, combing one's hair, putting trousers on, opening a door, talking on the telephone and drinking a glass of water). These instrumental displays elicit action representation (Johnson-Frey *et al.*, 2003) and are thus appropriate to use as controls for investigating emotional body expressions, also eliciting action representation and implicit movement perception.

Prior to the recordings the actors were briefed with a set of standardized instructions. For the instrumental body actions instructions specified the action to be performed. For emotional body actions instructions specified a familiar scenario (for example, opening a door and finding an armed robber in front of you). Static images were obtained from the videos by

selecting the most informative frame from the video file and converting it to grayscale pictures. To exclude that face recognition would play a role in recognition of the whole body stimuli, the faces were blocked with a grey mask. Stimulus selection for the present experiment was based on the results of a pilot study in which the images were presented one by one on a PC screen and shown for 4000ms with a 4000ms interval. For the emotional bodies, a total of 120 trials were used (3 expressions x 8 identities x 5 repetitions). Participants were instructed to categorize each stimulus in a forced-choice procedure as quickly and as accurately as possible by pressing one of the three response buttons corresponding to the three emotions. Overall correct recognition rate was between 100 and 65 percent (average 93%). For each stimulus category the six highest ranked exemplars were chosen (all recognized at 100% accuracy) for use in the present study. The same procedure was used for the instrumental gestures. Overall correct recognition rate in the pilot study was between 100 and 90 percent (average 99%), and for each gesture the six highest ranked were chosen (all recognized at 100% accuracy) for the present study.

The experiment consisted of randomized simultaneous presentation of three images on each trial, one at the top (target) and two probes right en left underneath. Target and probe pictures were always of three persons of the same sex, but one probe displayed the same expression as the target and the other one a different expression. Participants were required in a two alternative forced choice task to select the probe with the same expression as the target. They responded by pressing the corresponding button (see Fig 1A and 1B for stimulus examples). Stimuli were presented until response. During the intertrial interval (3000 ms) a blank screen was shown. The experiment consisted of a block with emotional stimuli and a block with instrumental stimuli.

The session started with four familiarization trials, followed by 75 experimental trials for the emotional block (3 emotions x 5 stimuli x 5 repetitions) and 48 trials for the instrumental block (6 actions x 4 stimuli x 2 repetitions). Similar to a previous study (Van den Stock et al. (2007), we used a matching task instead of a naming or categorization task because we were primarily interested in tapping into the motor perception processes. A naming or categorization task appeals more to verbal labeling of stimuli and also necessitates a significant language component, abilities which are not impaired in HD. The 2AFC task used here requires that matching be based on the basis of movement similarities between stimuli in the same emotion category.

Results

Mean accuracy scores for the two groups, separated by expression, are displayed in Figure 1C. We carried out a repeated measures ANOVA with expression (4 levels: instrumental, anger, fear and sadness) as within-subjects variable and group (2 levels: HD and Controls) as between-subjects variable. This revealed significant effects of type of expression, $F(3, 108) = 23.54, p < 0.001$; group, $F(1, 36) = 22.25, p < 0.001$; and a significant interaction, $F(3, 108) = 3.77, p = 0.013$. To follow up on the interaction effect, we performed for every expression Bonferroni corrected t -tests between both groups. This showed significant differences between the groups on the instrumental and angry expressions, $t(36) > 4.73, p < 0.001$. In order to explore differences between stage 1 and stage 2 HD, we performed a post hoc repeated measures ANOVA on the data of the patient group with expression (4 levels) as within-subjects variable and stage (2 levels) as between-subjects variable. This showed no main effect of group or interaction with expression.

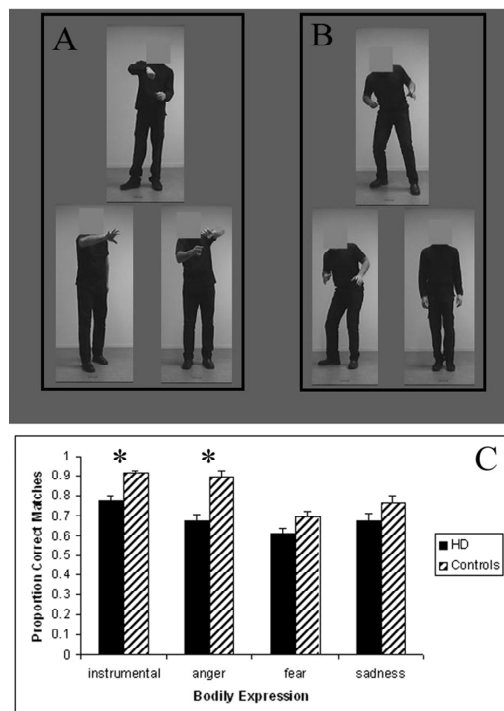


Figure 1. (A) Stimulus example in the instrumental action block and (B) the emotional block showing a fearful body on top and a fearful (left) and a sad (right) body at the bottom. (C) Proportion correct matches as a function of group and bodily expression. Error bars represent one SEM above the mean. * Significant at the $P < 0.01$ level.

To investigate a possible relationship between whole body recognition abilities on the one hand and behavioral and neurostructural deficits on the other hand, we performed partial correlation analyses, controlling for scores on the MDRS in order to correct for effects due to general cognitive decline.

Table 2 shows significant correlations between general assessment measures and experimental data. Partial correlations between bicaudate ratio and experimental data were not significant for any bodily expression. Since the UHDRS motor consists of a number of subscales, we performed an exploratory correlation analysis at the level of different components of the UHDRS motor, according to Shannon (Shannon, Raman, & Leurgans, 1999) namely chorea, dystonia, oculo-motor dysfunction and motor dysfunction. This revealed a significant correlation between the motor dysfunction component and recognition of angry body postures, $r = -0.575$, $p = 0.016$. The correlation between motor dysfunction and recognition of instrumental body postures was marginally significant, $r = -0.476$, $p = 0.053$. When confined to the subjects who underwent both bicaudate and UHDRS measurements ($N=9$), both correlations remained stable, but at a lower significance level ($r = -0.712$, $p = 0.048$ and $r = -0.643$, $p = 0.085$ respectively). Motor dysfunction takes into account gait disturbance, axial disorders, bradykinesia, rigidity, postural reflexes and gesture disabilities.

Table 2 *Partial Correlations (Controlling for Performance on MDRS) Between Bicaudate Ratio, UHDRS Subscales (Shannon et al., 1999) and Experimental Data.* * Significant at the $p < 0.05$ -level; (*) marginally significant

		bicaudate ratio	chorea	dystonia	motor	oculo-motor
body anger	<i>r</i>	-0,012	-0,046	-0,445	-0,575	-0,190
	<i>p</i>	0,973	0,861	0,073	0,016*	0,464
	N	11	15	15	15	15
body fear	<i>r</i>	0,111	0,123	0,298	0,006	-0,060
	<i>p</i>	0,761	0,640	0,245	0,983	0,819
	N	11	15	15	15	15
body sad	<i>r</i>	-0,555	0,085	-0,356	-0,182	-0,161

	<i>p</i>	0,096	0,747	0,161	0,484	0,536
	N	11	15	15	15	15
body instrumental	<i>r</i>	-0,231	0,001	-0,291	-0,476	-0,199
	<i>p</i>	0,520	0,996	0,257	0,053(*)	0,443
	N	11	15	15	15	15

Correlations between bicaudate ratio and MDRS, $r = -0.737$, $p = 0.004$ and between bicaudate ratio and the UHDRS motor dysfunction component, $r = 0.558$, $p = 0.047$ were also significant.

Discussion

The major finding of this study concerns the deficit of HD patients in recognizing instrumental and angry whole body postures. Recognition of meaningful non-emotional actions was not investigated previously, yet reports in the literature indicate that action related deficits in HD have been observed with different tasks in other settings (Aron, Sahakian, & Robbins, 2003). The present data provide evidence that action related abilities are important for recognition of instrumental actions and bodily expressions of anger. The action component at stake in recognition of whole body expressions of sadness is considerably less important. We conjecture that this is due to the fact that this emotion is typically associated with relaxation and loss of muscle tonus. Similarly, recognition of whole body expressions of fear also implies a reduced action component as fear cannot only lead to flight but is equally associated with freezing of the whole body (LeDoux, 1996). So the observed deficits in recognizing instrumental body actions and bodily expressions of anger are compatible with the idea that the motor deficit of HD patients impairs their ability for action recognition. On the basis of this, a relationship between the UHDRS motor score and experimental data was expected. This is in fact the case. Corrected for general cognitive decline, HD patients were more impaired in recognizing the emotional angry and the instrumental body expressions if their motor symptoms were more severe. The difficulties in perceiving body emotions were not related to abnormal movements like chorea or dystonia but to features that better capture the body posture and the abilities to perform gestures. This suggests a link between perception

and production of body gestures that needs to be further investigated. In view of our previous results that caudate nucleus activity was observed in the contrast between "fearful" and "instrumental" expressions of the body (de Gelder et al. 2004) it is surprising that neurodegeneration of the caudate nucleus in these HD patients is not reflected in impaired fear processing. But at present we have no data available directly comparing fear and anger expressions which would allow us to estimate the relative involvement of caudate nucleus in fear vs. anger action perception. This issue clearly needs further investigations.

The importance of motor areas for action recognition is illustrated by the research on mirror neurons by Rizzolatti and colleagues (di Pellegrino, Fadiga, Fogassi, Gallese, & Rizzolatti, 1992; Grèzes & Decety, 2002; Rizzolatti & Craighero, 2004). Research over the last decade has clearly indicated that recognition of instrumental actions involves some of the same brain areas that are involved in performance of that action by the observer himself. Thus the observed impairment in recognition of instrumental actions evokes the concept of motor resonance at the center of motor cognition abilities, which are implemented in premotor cortex, parietal cortex and superior temporal sulcus (STS). Degeneration of the motor areas in HD, predominantly striatum and its connections to parietal and premotor cortex and STS is consistent with the importance of action representation for intact recognition of whole body postures. The areas involved in spontaneous facial expressions (dorsolateral prefrontal cortex, orbitofrontal cortex, anterior cingulate, insula and amygdala) connect with the motor system via de basal ganglia (Alexander & Crutcher, 1990; Damasio, 1999) and this network may constitute an important part of a dedicated mechanism for visuomotor emotion perception.

Correlations between structural and functional cerebral changes and cognitive abilities have been reported in HD, although not consistently (see Montoya, Price, Menear, & Lepage, 2006, for a review). We were unable to find a correlation between our structural anatomical index (adjusted bicaudate ratio) and recognition of body postures. It should be stated that we preferred to include cortical pathology in our anatomical index, in view of recent reports about rather widespread cortical pathology already in preclinical HD (Kassubek, Bernhard Landwehrmeyer et al., 2004; Kassubek, Gaus et al., 2004; Thieben et al., 2002). However, incorporating cortical atrophy may rule out any kind of specific relationship between decreased volume of the striatum on the one hand and behavioral data on the other hand.

A question for future research concerns the relationship between recognition of emotional faces and bodies. Recently, selective deficits in recognition of angry faces have been reported in patients with damage to the ventral striatum (Calder, Keane, Lawrence, & Manes, 2004).

Furthermore, the disgust recognition deficit in HD has also been extended from facial expressions to scenes, odors, vocal expressions and declarative knowledge (Hayes, Stevenson, & Coltheart, 2007). We did not include whole bodies expressions of disgust, since our pilot data indicated they were very hard to recognize in static stimuli once the facial information is completely blurred.

Considering the relation we observed between emotion recognition deficits (specifically anger) and motor abilities, an interesting question is whether the same pattern of deficits observed here will also be found when we use dynamic stimuli. We are currently investigating this issue.

References

- Alexander, G. E., & Crutcher, M. D. (1990). Functional architecture of basal ganglia circuits: neural substrates of parallel processing. *Trends Neurosci*, 13(7), 266-271.
- Aron, A. R., Sahakian, B. J., & Robbins, T. W. (2003). Distractibility during selection-for-action: differential deficits in Huntington's disease and following frontal lobe damage. *Neuropsychologia*, 41(9), 1137-1147.
- Bhatia, K. P., & Marsden, C. D. (1994). The behavioural and motor consequences of focal lesions of the basal ganglia in man. *Brain*, 117 (Pt 4), 859-876.
- Calder, A. J., Keane, J., Lawrence, A. D., & Manes, F. (2004). Impaired recognition of anger following damage to the ventral striatum. *Brain*, 127(Pt 9), 1958-1969.
- Damasio, A. R. (1999). *The Feeling of What Happens*. New York: Harcourt Brace.
- de Gelder, B. (2006). Towards the neurobiology of emotional body language. *Nature Reviews Neuroscience*, 7(3), 242-249.
- de Gelder, B., Snyder, J., Greve, D., Gerard, G., & Hadjikhani, N. (2004). Fear fosters flight: A mechanism for fear contagion when perceiving emotion expressed by a whole body. *Proc Natl Acad Sci U S A*, 101(47), 16701-16706.
- di Pellegrino, G., Fadiga, L., Fogassi, L., Gallese, V., & Rizzolatti, G. (1992). Understanding motor events: a neurophysiological study. *Exp Brain Res.*, 91, 176-180.
- Grèzes, J., & Decety, J. (2002). Does visual perception of object afford action? Evidence from a neuroimaging study. *Neuropsychologia*, 40(2), 212-222.
- Grezes, J., Pichon, S., & de Gelder, B. (2007). Perceiving fear in dynamic body expressions. *Neuroimage*, 35(2), 959-967.
- Grillner, S., Hellgren, J., Menard, A., Saitoh, K., & Wikstrom, M. A. (2005). Mechanisms for selection of basic motor programs--roles for the striatum and pallidum. *Trends Neurosci*, 28(7), 364-370.
- Hayes, C. J., Stevenson, R. J., & Coltheart, M. (2007). Disgust and Huntington's disease. *Neuropsychologia*, 45(6), 1135-1151.

- Johnson-Frey, S. H., Maloof, F. R., Newman-Norlund, R., Farrer, C., Inati, S., & Grafton, S. T. (2003). Actions or hand-object interactions? Human inferior frontal cortex and action observation. *Neuron*, 39(6), 1053-1058.
- Kampe, K. K., Frith, C. D., Dolan, R. J., & Frith, U. (2001). Reward value of attractiveness and gaze. *Nature*, 413(6856), 589.
- Kassubek, J., Bernhard Landwehrmeyer, G., Ecker, D., Juengling, F. D., Mücke, R., Schuller, S., et al. (2004). Global cerebral atrophy in early stages of Huntington's disease: quantitative MRI study. *Neuroreport*, 15(2), 363-365.
- Kassubek, J., Gaus, W., & Landwehrmeyer, G. B. (2004). Evidence for more widespread cerebral pathology in early HD: an MRI-based morphometric analysis. *Neurology*, 62(3), 523-524; author reply 524.
- LeDoux, J. E. (1996). *The emotional brain: The mysterious underpinnings of emotional life*. New York, NY, US: Simon and Schuster.
- Mattis, S. (1976). Mental status examination for organic mental syndrome in elderly patients. In Bellak L, Karasu TB, editors. *Geriatric psychiatry*. New York: Grune & Straton.
- Montoya, A., Price, B. H., Menear, M., & Lepage, M. (2006). Brain imaging and cognitive dysfunctions in Huntington's disease. *J Psychiatry Neurosci*, 31(1), 21-29.
- Rizzolatti, G., & Craighero, L. (2004). The mirror-neuron system. *Annu Rev Neurosci*, 27, 169-192.
- Shannon, K. M., Raman, R., & Leurgans, S. (1999). Primary clinical components of Huntington's disease: factor structure of the motor subscale of the Unified Huntington's Disease Rating Scale. *Neurology*, 52((suppl 2)), A120.
- Shoulson, I. (1981). Huntington disease: functional capacities in patients treated with neuroleptic and antidepressant drugs. *Neurology*, 31(10), 1333-1335.
- Sprengelmeyer, R., Young, A. W., Calder, A. J., Karnat, A., Lange, H., Homberg, V., et al. (1996). Loss of disgust. Perception of faces and emotions in Huntington's disease. *Brain*, 119 (Pt 5), 1647-1665.
- Thieben, M. J., Duggins, A. J., Good, C. D., Gomes, L., Mahant, N., Richards, F., et al. (2002). The distribution of structural neuropathology in pre-clinical Huntington's disease. *Brain*, 125(Pt 8), 1815-1828.

- Van den Stock, J., Righart, R., & de Gelder, B. (2007). Body expressions influence recognition of emotions in the face and voice. *Emotion*, 7(3), 487-494.
- Vonsattel, J. P., Myers, R. H., Stevens, T. J., Ferrante, R. J., Bird, E. D., & Richardson, E. P., Jr. (1985). Neuropathological classification of Huntington's disease. *J Neuropathol Exp Neurol*, 44(6), 559-577.
- Wang, K., Hoosain, R., Yang, R. M., Meng, Y., & Wang, C. Q. (2003). Impairment of recognition of disgust in Chinese with Huntington's or Wilson's disease. *Neuropsychologia*, 41(5), 527-537.

Chapter 11: Perceiving emotions from bodily expressions and multisensory integration of emotion cues in schizophrenia¹¹

Introduction

Investigations of emotion recognition in schizophrenics have predominantly focussed on facial expressions. The findings point to a deficit in recognition of negative emotions (Mandal, Pandey, & Prasad, 1998). This facial emotion recognition impairment has been linked to the social dysfunctions observed in schizophrenics (Pinkham, Hopfinger, Ruparel, & Penn, 2008). From that perspective a facial expression deficit is not all that surprising. But an important issue is whether one can generalize from a deficit in recognition of facial expressions to difficulties in recognizing emotional signals conveyed by other common channels like the voice and the body. In daily life, emotions are also expressed in the whole body and by different sensory channels, but studies addressing body language recognition or emotional multisensory integration in schizophrenia so far are rare. We recently used affective face-voice combinations that were either congruent (for example a happy face presented simultaneously with a happy vocal expression) or incongruent (for example a happy face paired with a fearful vocal expression) and asked schizophrenics to rate one of both modalities and ignore the other (de Gelder et al., 2005; de Jong, Houdiamont, Van den Stock, & de Gelder, 2009). The results showed that schizophrenics showed anomalous crossmodal bias effects, compared to control subjects. For example, when schizophrenic patients categorized the emotion expressed in the voice, they were less influenced by the simultaneously presented but task irrelevant facial expression (de Jong et al., 2009). So far, little is known about recognition of emotional body language in schizophrenia, but in view of the behavioural and neuro-anatomical similarities between perception of faces and bodies (de Gelder, 2006; de Gelder et al., 2009), we hypothesize that the patients will be impaired.

Our goal here was to explore these two issues further. In Experiment 1, we investigated the recognition of emotional body language in a group of schizophrenics, non-schizophrenic psychotics and normal controls in order to explore whether the emotion recognition deficit previously reported for faces also applies for whole body expressions.

¹¹ This chapter is submitted

In Experiment 2 we investigated how schizophrenics perceive multisensory emotional events, consisting of realistic body language combined with affective auditory utterances. We presented video clips of emotional body language, engaged in a common activity in an everyday situation. In addition to adding human vocal expressions, we also combined the video clips with animal vocalizations in order to investigate the role of environmental sounds. As reported previously, it is important to control for task variables as attention may shift across conditions and trials from face to voice especially in clinical populations (Bertelson & de Gelder, 2004; de Gelder & Bertelson, 2003).

Experiment 1: Recognition of static body language

Methods

Participants. Thirty-one schizophrenics, 23 patients with non-schizophrenic psychosis and a group of 21 normal controls matched for gender, age and socio-economic status participated in the study. Demographic data are shown in Table 1. There was no significant difference in age ($F(2,72) = 0.620$; $p < 0.541$) or gender ratio ($\chi^2 \leq 2.13$; $p < 0.14$) between the three groups. Only patients meeting the criteria for schizophrenia and non-schizophrenic psychosis set by the DSM-IV (APA, 2000) were included. All patients were under treatment at the local day hospital. Diagnosis was established with the Schedules for Clinical Assessment in Neuropsychiatry (SCAN, version 2.1), a standardised interview for diagnosing axis I disorders, conducted by a trained psychiatrist. All patients were on antipsychotic medication. Control subjects with a psychiatric disorder, a brain dysfunction or a genetic predisposition for schizophrenia were excluded from participation. All participants were paid for participation (22€).

Informed consent was obtained from all participants according to the declaration of Helsinki and the study was approved by the regional Medical Ethics Committee.

Table 1. Demographic data

	Schizophrenia group	Non-schizophrenic psychosis group	Control group
Demographic data			
Experiment 1			
N	31	23	21
Mean age (range)	33.7 (21-52)	35.7 (20-54)	32.4 (21-58)
Gender	23 M/ 8 F	14 M/ 9 F	13 M/ 8 F
Dexterity	28 R/ 3 L	20 R/ 3 L	20 R/ 1 L
Experiment 2			
N	16	/	16
Mean age (range)	36.8 (22-53)	/	38.0 (22-53)
Gender	15 M/ 1 F	/	9 M/ 7 F
Dexterity	15 R/ 1 L	/	13 R/ 3 L

Materials and procedure

Materials consisted of pictures from our own database of body expressions and instrumental actions ((for details on stimulus construction, see refs de Gelder, Snyder, Greve, Gerard, & Hadjikhani, 2004; Hadjikhani & de Gelder, 2003)). All body images were shown with the faces blurred and had been validated in a previous study(Hadjikhani & de Gelder, 2003). Only pictures that were correctly recognized above 85 percent were selected for the experiment.

The experiment consisted of two blocks: one with bodily expressions and one with neutral bodily actions. We included these instrumental whole body actions, because these displays elicit action representation(Johnson-Frey et al., 2003) and are thus appropriate to use as controls for investigating emotional body expressions, also eliciting action representation and implicit movement perception. The procedure was identical in each block. Materials for the experiment consisted of 30 emotional bodies and 24 instrumental actions. The instrumental action block consisted of 48 trials (6 actions x 2 genders x 4 exemplars) and the bodily expression block consisted of 36 trials (3 expressions x 2 genders x 6 exemplars). A stimulus consisted of presentation of a target at the top of the screen that had to be matched with one of two simultaneously presented probes underneath (see Figure 1). The three pictures in a stimulus were always from the same gender but three different identities. Participants were instructed to select the probe that matched the action or emotion of the target body. The position of the correct probe was counterbalanced. Participants responded by pressing the corresponding button, indicating their choice for the left or right probe. Stimuli

were presented till a response was given. During the 1000 ms intertrial interval, a blank screen was shown.

Results

Results are displayed in Figure 1. We calculated for every condition and participant the mean accuracy and median reaction times (RT) of the correct trials. Both RT and accuracy data were submitted to a repeated measures ANOVA with expression (4 levels: angry, fearful, sad and instrumental) as within subjects factor and group (3 levels: schizophrenic, non-schizophrenic psychosis and control) as between subjects factor. This revealed for the accuracy data a significant main effect of expression ($F(3,210) = 13.269$; $p < 0.001$) and group ($F(2,70) = 6.234$; $p < 0.003$). The expression \times group interaction was not significant. Tukey post hoc tests on the main effect of group showed a significant difference between the control group and the schizophrenic group ($p < 0.002$), and a marginally significant difference between the control group and the non-schizophrenic psychosis group ($p < 0.054$). To follow up on the main effect of expression, we performed Bonferroni corrected paired sample t-tests between every combination of expressions ($n=6$). This showed significant differences between angry and fearful ($t(74) = 5.911$; $p < 0.001$); between instrumental and fearful ($t(74) = 6.818$; $p < 0.001$); and between instrumental and sad expressions ($t(74) = 4.303$; $p < 0.001$). The difference between fearful and sad expressions was marginally significant ($t(74) = 2.483$; $p < 0.015$).

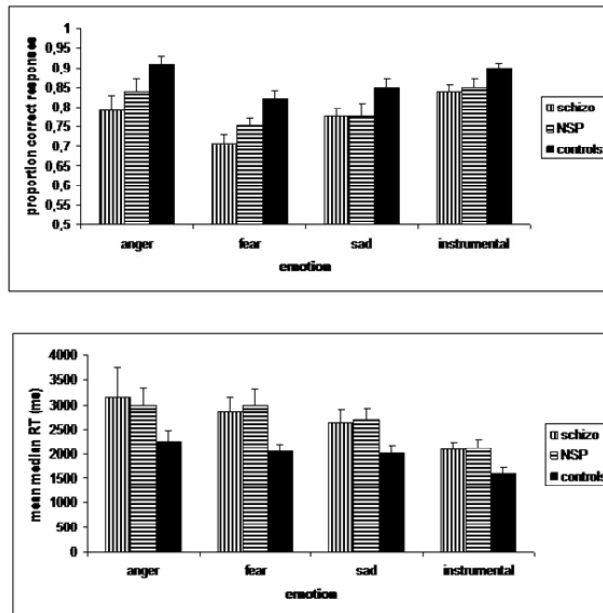


Figure 1. Accuracy (top) and reaction time (bottom) of Experiment 1 as a function of expression and group (Schizo = Schizophrenia group; NSP : non-schizophrenic psychosis group). Error bars represent 1 SEM.

The analysis on RT showed a significant main effect of expression ($F(3,210) = 8.762$; $p < 0.003$). The main effect of group and the expression \times group interaction were not significant. Bonferroni corrected paired sampled post hoc t-tests showed significant difference between angry and instrumental ($t(74) = 3.481$; $p < 0.001$); fearful and sad ($t(74) = 3.326$; $p < 0.001$); fearful and instrumental ($t(74) = 5.142$; $p < 0.001$); and between sad and instrumental ($t(74) = 4.009$; $p < 0.001$) expressions.

Discussion

We presented patients with schizophrenia, patients with non-schizophrenic psychosis and matched controls with a two alternative forced choice whole body expression matching task. The results show that compared to the control group, the schizophrenic group exhibits a general impairment in recognizing emotional body language. The non-schizophrenic psychosis group occupies an intermediate position between the controls and schizophrenia group. The impairment is manifest in both accuracy and reaction time data. The differences

between emotions in the accuracy and reaction time data are in the opposite direction and hence argue against speed-accuracy tradeoffs. It is also unlikely that the observed effects can be explained by task difficulty, since the absence of an expression x group interaction reveals that the patients are not differentially impaired on recognition of specific emotions, while the main effect of emotion indicates that not all emotions are equally well recognizable. Hence, the patients are not more impaired in recognizing the more difficult emotions.

The generalized whole body emotion recognition deficit is consistent with findings from facial expression recognition studies (reviewed in ref Mandal et al., 1998). Schizophrenics are in general less able to make adequate emotional judgements of ambiguous facial expressions (Kee, Horan, Wynn, Mintz, & Green, 2006) and attribute negative emotional valence to neutral face cues (Kohler et al., 2003), which may explain the non-specific nature of emotion recognition difficulties.

Experiment 2: Multisensory integration of dynamic body language and vocalizations

In everyday situations, fearful body language is usually accompanied by anxious screams. Recently, we showed that static whole body expressions influence recognition of simultaneously presented vocal expressions (Van den Stock, Righart, & de Gelder, 2007). In a follow-up study (Van den Stock, Grezes, & de Gelder, 2008) we used dynamic stimuli in realistic situations to increase ecological validity, which may be an important factor in multisensory integration (Bertelson & de Gelder, 2004; de Gelder & Bertelson, 2003). We paired these visual stimuli with nonverbal vocalizations and more importantly, we also manipulated the nature of the bimodal combinations by using auditory stimuli that were either produced by the same source as the visual stimuli (human vocalizations), or by a different source (animal vocalizations). The findings showed that both human and animal sounds influence recognition of dynamic body language. Here we use the same materials for the second objective of the present study which is to investigate the multisensory integration pattern of these everyday emotional events in schizophrenia.

Methods

Participants. A new group of sixteen schizophrenics meeting the criteria described in Experiment 1 and sixteen matched controls participated in Experiment 2. Demographic data are shown in Table 1. There was no significant difference of age between groups ($t(30) = 0.32, p < 0.751$). None of the participants of Experiment 2 participated in Experiment 1.

Materials and procedure

Materials and procedure of experiment 2 are described in detail in ref(2008). In summary, the experiment consisted of a visual (V) block, in which video clips of an actor grabbing a glass in a fearful or happy manner were presented and audio-visual (AV) block in which each video was synchronically presented with one of the stimuli of the four auditory conditions: fearful human vocalizations, happy human vocalizations, fearful animal vocalizations and happy animal vocalizations. Participants were instructed to categorize the emotion expressed by the body while ignoring the auditory message in a two alternative forced choice task by pressing the corresponding button (happy or fearful).

Results

We excluded trials on which participants responded before the end of the stimulus ($RT < 800ms$). On this basis, sixty-four trials (1.3%) were discarded. We computed the proportion 'happy' responses in the different conditions. Results are shown in Figure 2. The data with animal and human vocalizations are analyzed separately. Since the participants performed a delayed reaction time task, RT data were not analyzed.

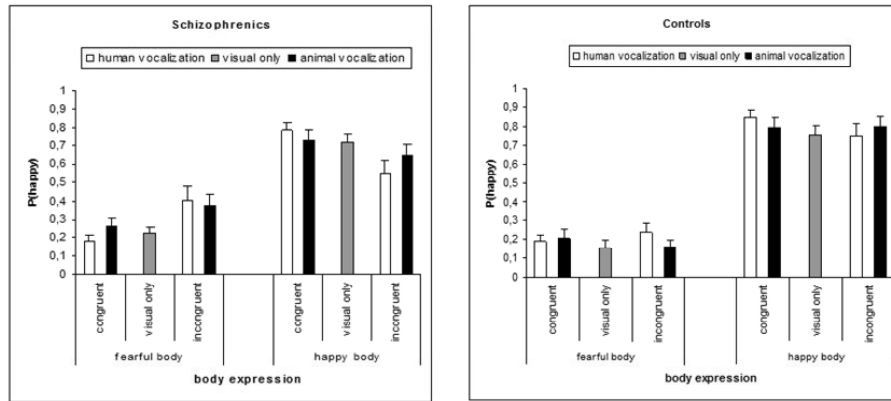


Figure 2. Proportion 'happy' responses in the bimodal and unimodal conditions, separated by group, emotion, auditory category and congruence. Error bars represent 1 SEM.

a) body videos simultaneously presented with human vocalizations:

A repeated measures ANOVA was performed on the proportion 'happy' responses with visual emotion (two levels: fearful and happy) and (human) auditory emotion (three levels: fearful, happy and no auditory stimulus) as within-subjects factors and with group (two levels: schizophrenia and control) as between subjects factor. This revealed a main effect of visual emotion ($F(1,31) = 124.154$, $p < 0.001$), a main effect of auditory emotion ($F(1,31) = 11.278$, $p < 0.001$) and a significant two way auditory emotion \times group interaction ($F(2,62) = 3.310$, $p < .043$). The visual emotion \times group interaction was marginally significant ($F(1,31) = 3.937$, $p < 0.056$).

The main effect of visual emotion indicates, as expected that the proportion happy responses is higher for happy bodies than for fearful bodies. The main effect of auditory emotion shows that the ratings of the bodily expressions are influenced by the auditory emotion, while the auditory emotion \times group interaction indicates that this auditory influence is significantly different between the two groups. To follow up on this interaction effect we computed the influence of the auditory information for both groups separately, by calculating the ordinal difference between the unimodal and bimodal conditions [(fear video minus fear video paired with fear audio) + (fear video paired with happy audio minus fear video) + (happy video paired with happy audio minus happy video) + (happy video minus happy video paired with fearful audio)]. The resulting difference was higher for the schizophrenia group

(0.46) compared to the control group (0.14) ($t(31) < 2.036$, $p < 0.054$), indicating the schizophrenics are more influenced by the vocalizations than the controls.

We also compared the ratings of both unimodal visual conditions (fearful and happy body language) between both groups using independent samples t -tests. This showed no significant difference ($t(31) < 1.335$, $p < 0.192$), indicating both patients and controls were equally able to recognize the body videos.

b) Body videos simultaneously presented with animal vocalizations:

A repeated measures ANOVA on the proportion happy responses with visual emotion (fearful and happy) and (animal) auditory emotion (fearful, happy and no auditory stimulus) as within-subjects factors, and group (schizophrenia and control) as between subjects factor, revealed a significant main effect of visual emotion ($F(1,31) = 112.758$, $p < 0.001$) and a significant visual emotion x group interaction ($F(1,31) = 4.456$, $p < 0.043$). To follow up on the interaction effect, we computed for both groups separately, the mean proportion happy responses for the conditions with a happy video, regardless of (animal) auditory information and we followed the same procedure for the fearful video conditions. The proportion happy responses on the conditions with fear videos was significantly lower in the control group than in the schizophrenia group ($t(31) = 2.199$, $p < 0.035$), suggesting that the schizophrenics categorized the fearful videos more as happy. The difference between both groups on the conditions with happy videos was not significant.

A comparison between both groups on recognition of each of the four unimodal auditory conditions showed no significant difference ($t(31) < 1.850$, $p < 0.074$).

Discussion

We presented schizophrenics and controls with short videos of a person engaged in an everyday action (picking up a glass from the table), performed either in a fearful or happy manner and asked them to indicate the emotion displayed by the actor. Simultaneously with the videos we presented emotionally congruent or incongruent vocal expressions, which could be produced by a human or an animal. The results show that schizophrenics are more influenced by the task irrelevant auditory information, compared to the control group, but only for human and not for the animal vocalizations. Different explanations may be envisaged for the increased influence of human auditory information in the schizophrenia

group. One may relate to the relative dominance of sensory inputs. The increased crossmodal bias of vocal expressions on body expressions may point to greater impact of the auditory modality under audiovisual perception conditions in schizophrenics. This explanation is also compatible with our previous study in which schizophrenia patients showed a reduced crossmodal bias of a visual facial expression on the recognition of the emotion in a vocal expression (de Jong et al., 2009) and with a recent report from audiovisual speech perception in schizophrenia (Ross et al., 2007). But this does not explain why human and animal vocalisations have a different effect. Another possible explanation may be task difficulty. The recognition of the unimodal visual stimuli shows no ceiling effect and if schizophrenics have more difficulty recognizing the visual stimulus, they might rely more on the information provided by the secondary stimulus. A direct test between both groups of the unimodal conditions reveals no significant difference, indicating that schizophrenics and controls perform equally in recognizing whole body expressions as well as human and auditory vocalizations, but this does not rule out the possibility that they have more difficulty with audiovisual stimuli since performance on an audiovisual integration task depends on other factors besides performance in the respective sensory systems. The auditory information may be harder to ignore for the schizophrenics either because the focussed attention task requiring to shut out one input system or because it is harder in the case of human sounds.

General discussion

In the first experiment, we tested recognition of static emotional body language in a group of schizophrenics, non-schizophrenic psychotics and controls. The results show an emotion recognition impairment in the schizophrenic group for all emotions. The impairment is also present in the non-schizophrenic psychosis group, but to a lesser extent. The present study shows that the emotion recognition difficulties in schizophrenia, which have been previously documented with studies using facial expressions (Mandal et al., 1998), extend to the recognition of body language. These perceptual deficits are compatible with neuro-anatomical findings: perception of bodily expressions activates not only brain areas associated with emotion perception like fusiform gyrus and amygdala (Hadjikhani & de Gelder, 2003), but also areas involved in action representation like premotor cortex (de Gelder et al., 2004; Grèzes, Pichon, & de Gelder, 2007; Pichon, de Gelder, & Grèzes, 2008) and all these

structures show abnormalities in schizophrenia (Bertrand et al., 2008; Gur et al., 2002; Michalopoulou et al., 2008; Phillips et al., 1999; Pinkham, Hopfinger, Pelphrey, Piven, & Penn, 2008; Pinkham, Hopfinger, Ruparel et al., 2008). So impaired recognition of body expressions in schizophrenia might have its roots in a dysfunction of the brain network involved in emotion perception, but also in a deficit of the brain areas involved in action representation.

It has recently been suggested that the motor abilities of the observer are an important aspect of body language recognition, linking movement deficits and anomalous recognition of bodily expressions (de Gelder, 2006; de Gelder, Van den Stock, de Diego Balaguer, & Bachoud-Levi, 2008). It is possible that the motor problems associated with schizophrenia, like catatonia, play an important role in recognizing emotional body language and this may be an interesting question to explore in future studies.

The second experiment focussed on multisensory integration of dynamic emotional body language on the one hand and both human and animal vocalizations on the other hand. The data show an increased integration of both modalities in the schizophrenic group, but only when the auditory information consists of human voices, suggesting that ecological validity plays a role in the deficient multisensory integration. We recently reported a reduced crossmodal bias effect of emotional faces on the recognition of the emotion in a vocal expression (de Jong et al., 2009). Both these data sets are compatible with an auditory dominance in the schizophrenia group when perceiving audiovisual stimuli of which both modalities have a high rate of co-occurrence in daily life. Compatible with this notion, a recent study showed that schizophrenia patients were less able than controls to benefit from visual speech information when recognizing degraded auditory speech (Ross et al., 2007).

At the neuro-anatomical level, binding of emotional information in the face and voice has been associated with activity in the amygdala (Dolan, Morris, & de Gelder, 2001; Ethofer et al., 2006) and abnormal amygdala activity has been reported in schizophrenics in response to facial expressions (Gur et al., 2002; Michalopoulou et al., 2008; Phillips et al., 1999). It is therefore not unlikely that the anomalous multisensory integration may partly have its roots in abnormal amygdala activity.

The importance of the amygdala as discussed in both experiments is not surprisingly considering its established involvement in emotion perception (Zald, 2003), but another important neuro-anatomical aspect relates to brain connectivity. A recent fMRI study reported

anomalous connectivity between amygdala and frontal regions in a group of schizophrenia patients (Leitman et al., 2008). Abnormal amygdalar-frontal connectivity may either be cause or effect of dysfunctional amygdala and combining both findings provides a neuro-anatomical basis for both the emotion recognition deficit (Experiment 1) and the anomalous multisensory integration (Experiment 2).

Interestingly, the patients show a general impairment in recognizing static emotional body expressions, whereas with the videos used here there is no significant group difference. Obviously, videos present the viewer with more information than still pictures and are therefore relatively easier to recognize. Next to the methodological differences, a possible explanation for this finding concerns the cognitive demands. Recognizing isolated static expressions requires the perceptual system to fill in spontaneously the missing information or requires the brain to engage in motor imagery in order to compensate for the lack of physical information in the stimulus, like direction of movement and speed of movement. Future research is needed to identify the specific processes that are impaired in schizophrenics, when recognizing affective stimuli, possibly related to ecological validity and motor symptoms.

References

- APA. (2000). *Diagnostic and Statistical Manual of Mental Disorders, DSM-IV-TR*. Washington, DC: American Psychiatric Association.
- Bertelson, P., & de Gelder, B. (2004). The psychology of multimodal perception. In C. Spence & J. Driver (Eds.), *Crossmodal space and crossmodal attention* (pp. 151-177). Oxford: Oxford University Press.
- Bertrand, M. C., Achim, A. M., Harvey, P. O., Sutton, H., Malla, A. K., & Lepage, M. (2008). Structural neural correlates of impairments in social cognition in first episode psychosis. *Soc Neurosci*, 3(1), 79-88.
- de Gelder, B. (2006). Towards the neurobiology of emotional body language. *Nature Reviews Neuroscience*, 7(3), 242-249.
- de Gelder, B., & Bertelson, P. (2003). Multisensory integration, perception and ecological validity. *Trends Cogn Sci*, 7(10), 460-467.
- de Gelder, B., Snyder, J., Greve, D., Gerard, G., & Hadjikhani, N. (2004). Fear fosters flight: A mechanism for fear contagion when perceiving emotion expressed by a whole body. *Proc Natl Acad Sci U S A*, 101(47), 16701-16706.
- de Gelder, B., Van den Stock, J., de Diego Balaguer, R., & Bachoud-Levi, A. C. (2008). Huntington's disease impairs recognition of angry and instrumental body language. *Neuropsychologia*, 46(1), 369-373.
- de Gelder, B., Van den Stock, J., Meeren, H. K., Sinke, C. B., Kret, M. E., & Tamietto, M. (2009). Standing up for the body. Recent progress in uncovering the networks involved in processing bodies and bodily expressions. *Neuroscience and Biobehavioral Reviews*.
- de Gelder, B., Vroomen, J., de Jong, S. J., Masthoff, E. D., Trompenaars, F. J., & Hodiament, P. (2005). Multisensory integration of emotional faces and voices in schizophrenics. *Schizophr Res*, 72(2-3), 195-203.
- de Jong, J. J., Hodiament, P. P., Van den Stock, J., & de Gelder, B. (2009). Audiovisual emotion recognition in schizophrenia: reduced integration of facial and vocal affect. *Schizophrenia Research*, 107(2-3), 286-293.
- Dolan, R. J., Morris, J. S., & de Gelder, B. (2001). Crossmodal binding of fear in voice and face. *Proc Natl Acad Sci U S A*, 98(17), 10006-10010.

- Ethofer, T., Anders, S., Erb, M., Droll, C., Royen, L., Saur, R., et al. (2006). Impact of voice on emotional judgment of faces: an event-related fMRI study. *Hum Brain Mapp*, 27(9), 707-714.
- Grèzes, J., Pichon, S., & de Gelder, B. (2007). Perceiving fear in dynamic body expressions. *Neuroimage*, 35(2), 959-967.
- Gur, R. E., McGrath, C., Chan, R. M., Schroeder, L., Turner, T., Turetsky, B. I., et al. (2002). An fMRI study of facial emotion processing in patients with schizophrenia. *Am J Psychiatry*, 159(12), 1992-1999.
- Hadjikhani, N., & de Gelder, B. (2003). Seeing fearful body expressions activates the fusiform cortex and amygdala. *Curr Biol*, 13(24), 2201-2205.
- Johnson-Frey, S. H., Maloof, F. R., Newman-Norlund, R., Farrer, C., Inati, S., & Grafton, S. T. (2003). Actions or hand-object interactions? Human inferior frontal cortex and action observation. *Neuron*, 39(6), 1053-1058.
- Kee, K. S., Horan, W. P., Wynn, J. K., Mintz, J., & Green, M. F. (2006). An analysis of categorical perception of facial emotion in schizophrenia. *Schizophr Res*, 87(1-3), 228-237.
- Kohler, C. G., Turner, T. H., Bilker, W. B., Brensinger, C. M., Siegel, S. J., Kanes, S. J., et al. (2003). Facial emotion recognition in schizophrenia: intensity effects and error pattern. *Am J Psychiatry*, 160(10), 1768-1774.
- Leitman, D. I., Loughhead, J., Wolf, D. H., Ruparel, K., Kohler, C. G., Elliott, M. A., et al. (2008). Abnormal superior temporal connectivity during fear perception in schizophrenia. *Schizophr Bull*, 34(4), 673-678.
- Mandal, M. K., Pandey, R., & Prasad, A. B. (1998). Facial expressions of emotions and schizophrenia: a review. *Schizophr Bull*, 24(3), 399-412.
- Michalopoulou, P. G., Surguladze, S., Morley, L. A., Giampietro, V. P., Murray, R. M., & Shergill, S. S. (2008). Facial fear processing and psychotic symptoms in schizophrenia: functional magnetic resonance imaging study. *Br J Psychiatry*, 192(3), 191-196.
- Phillips, M. L., Williams, L., Senior, C., Bullmore, E. T., Brammer, M. J., Andrew, C., et al. (1999). A differential neural response to threatening and non-threatening negative facial expressions in paranoid and non-paranoid schizophrenics. *Psychiatry Res*, 92(1), 11-31.
- Pichon, S., de Gelder, B., & Grèzes, J. (2008). Emotional modulation of visual and motor areas by dynamic body expressions of anger. *Social Neuroscience*, 3(3-4), 199-212.

- Pinkham, A. E., Hopfinger, J. B., Pelphrey, K. A., Piven, J., & Penn, D. L. (2008). Neural bases for impaired social cognition in schizophrenia and autism spectrum disorders. *Schizophr Res*, 99(1-3), 164-175.
- Pinkham, A. E., Hopfinger, J. B., Ruparel, K., & Penn, D. L. (2008). An investigation of the relationship between activation of a social cognitive neural network and social functioning. *Schizophr Bull*, 34(4), 688-697.
- Ross, L. A., Saint-Amour, D., Leavitt, V. M., Molholm, S., Javitt, D. C., & Foxe, J. J. (2007). Impaired multisensory processing in schizophrenia: deficits in the visual enhancement of speech comprehension under noisy environmental conditions. *Schizophr Res*, 97(1-3), 173-183.
- Van den Stock, J., Grezes, J., & de Gelder, B. (2008). Human and animal sounds influence recognition of body language. *Brain Research*, 1242, 185-190.
- Van den Stock, J., Righart, R., & de Gelder, B. (2007). Body expressions influence recognition of emotions in the face and voice. *Emotion*, 7(3), 487-494.
- Zald, D. H. (2003). The human amygdala and the emotional evaluation of sensory stimuli. *Brain Res Brain Res Rev*, 41(1), 88-123.

Discussion

The social brain hypothesis

Social animals have developed a whole range of communicative abilities on which their well-being and survival in a society hinges. Among the cognitive skills of social species is the ability to monitor each others' behavior and to adapt continuously to the social signals of others, whether collaborative or competitive most pronounced (Darwin 1872). A longstanding assumption in the literature is that dedicated brain resources exist devoted to processing social signals, a 'social brain' for short, and that social skills resulted from brain evolution in a social context. Evidence to support this assumption is intended to apply across widely different social species (Preston & de Waal, 2002), all of which have to rely upon conspecifics in order to survive.

The hypothesis of a social brain encompassing specialized skills for social perception and cognition has gained in popularity among neuroscientists over the last decade. Supportive findings are gathered from a wide spectrum of disciplines, ranging from behavioral observations to single cell recordings. Species with a similar social history can be expected to exhibit important similarities in the lay-out of a 'social brain'. This assumption is the backbone of much neurobiological emotion research of the last two decades where results from animal research have inspired human studies (Adolphs, 2002; R.J. Dolan, 2002; Phelps & Ledoux, 2005).

Traditionally the social brain hypothesis is tested in the visual modality as most of the primate cortex is dedicated to visual processing and it makes sense to assume that the visual system is particularly attuned to the perception of social cues. The visual objects to which the social brain is likely to be sensitive comprises facial expressions and whole body gestures, postures and movements. They all provide salient cues for adaptive behavior, as seen very clearly in displays of aggression which automatically function as a warning signal.

The main topic of this thesis focuses on how our perception of whole body expressions is not solely shaped by the body itself, but also by contextual stimuli like auditory information and scenes. The subject of this thesis itself is somewhat illustrative for the evolution of the field of

human affective neuroscience during the last decade. Until recently, the scientific knowledge of emotion perception was based almost exclusively on facial expression research. Moreover, the literature contained a significant bias since the bulk of the studies used the stimulus set developed by Ekman in the seventies (Ekman & Friesen, 1976). This thesis comprises a double shift from the isolated facial expression perception inheritance.

Emotions are expressed in the whole body

First, the spotlight on facial expressions zoomed out to shed light on whole body expressions. It is somewhat surprising that the detailed description of bodily expressions of many different emotions already by Darwin in the nineteenth century (1872), was not followed upon more systematically. With the exception of a few isolated reports (Argyle, 1988; Sprengelmeyer et al., 1999), the literature on how body expressions are processed has only taken off in the last decade. One of the first basic research questions concerned whether observers can easily recognize different emotional states based on body expressions alone. The available data indicate that this is clearly the case (Van den Stock, Righart, & de Gelder, 2007). This is not surprisingly considering the high frequency of interactions with conspecifics. Repeated exposure to body language, be it emotional or neutral, leads to perceptual expertise and tuning of the visual system. A similar explanation has been proposed for perception of facial expressions (e.g. Gauthier, Skudlarski, Gore, & Anderson, 2000).

When investigating emotional body language, comparing the results with what is known from facial expression research is almost inevitable, considering the many similarities between both stimulus categories. Bodies and faces both provide information on diverse dimensions like identity, emotion, gender, age... Behavioral, neuro-anatomical and neuro-temporal similarities between bodies and faces are reviewed in (B. de Gelder, 2006; B. de Gelder et al., 2009; Peelen & Downing, 2007). An interesting approach in comparing findings from face and body research might be to focus on the differences rather than on the similarities. At face value, at least two significant differences between faces and bodies pop out.

faces convey information about identity

First, faces provide significantly more information about identity compared to bodies. Headless bodies reveal little information about personal identity, whereas faces alone are sufficient for person identification. The fact that bodies contain little identity information is related to the fact that bodies are usually clothed. Clothing may conceal bodily features that are sufficient for identification. On the other hand, it has been shown that people can recognize friends by dynamic information provided by the body alone (Cutting & Kozlowski, 1977), underscoring the importance of dynamic and personal identity information contained by body expressions.

bodies convey information about actions

Secondly, the emphasis on the function of facial expressions lies in communication of the emotional state of the producer, whereas whole body expressions also serve adaptive behavioral functions, like fight or flight. There is little discussion that the behavioral phenotype of facial expressions is associated with adaptive functions, but these lay primarily in the social domain, for example cooperative interaction signaling or infant-caregiver interaction facilitation (Schmidt & Cohn, 2001). Another important function of facial expressions is danger signaling, where gaze direction also communicates the location of the threat. Traditionally, facial expression research did not target the adaptiveness or communicative value of facial expressions, but started from the putative universal categorization of facial expression classes: anger, disgust, fear, sad, surprise and happy. As stated above, applicable methodology consisted of presentation of isolated faces, whereas variability in behavioral phenotypes of a particular emotional face expression and its relevance to the underlying adaptive function is considered as noise. Little consideration has been applied to the notion that many different kinds of for example fear exist. It is easy to imagine that fear of the dark, stage fright, separation anxiety and fear from an armed robber are associated with different facial expressions and adaptive behavior, although all can be labeled as 'fearful'. The adaptive behavior associated with the different events described above is more explicitly pronounced in the emotional expression conveyed by the whole body. Separation anxiety will result in approach behavior towards the object of affection, whereas confrontation with an armed robber will lead to avoidance tendencies. This means the communicative value of whole body expressions extends beyond the emotional state of

the producer, but also includes information about the action tendencies of the producer. This is an important implication and exposes the direct link between emotion and action, a link that hidden in facial expressions.

How actions are represented in the brain has been a hot topic on the scientific research agenda, following the discovery of the currently famous mirror neurons. Mirror neurons were discovered in the parietal and frontal cortex of monkeys in neurophysiological studies and these neurons fire both when a monkey performs an action (for example grasping an object) and when a monkey sees the action being performed. A review of the mirror neuron system can be found in (Rizzolatti & Craighero, 2004). The focus on research on the mirror-neuron system has not been on its role in emotion perception, but it nevertheless brings forward interesting challenges. Specifically, the brain activation that has been repeatedly reported in human premotor cortex when observing emotional body language (B. de Gelder, Snyder, Greve, Gerard, & Hadjikhani, 2004; Grèzes, Pichon, & de Gelder, 2007; Pichon, de Gelder, & Grèzes, 2008) puts forward the question whether this reflects mere imitational processes rather than emotional contagion or action potentiation. Evidence supporting the latter hypothesis is available from studies reporting more premotor activity for emotional than for neutral bodies (B. de Gelder, Snyder et al., 2004; Grèzes et al., 2007; Pichon et al., 2008).

Other important advantages of studying whole body expressions in addition to facial expressions can be found in (B de Gelder, 2009).

Emotional expressions are prone to contextual influences

The second main turn from the isolated facial expression research concerns the investigation of the malleability of emotional expressions by contextual influences. While the emphasis used to lay on perception of isolated stimuli, stripped of any contextual information, the studies reported in this thesis focus on how our perception of well defined stimulus categories like bodily expressions are also shaped by simultaneously occurring visual and auditory information. This approach is more in line with our perceptions of everyday life which is dynamic and consists of a rich pallet of multimodal stimuli. Our research approach, moving beyond isolated stimulus categories therefore also reflects an increase in ecological validity. Our behavior and brains are tuned to perform optimally in our natural environment, and a

better understanding of the working of the brain and the laws of behavior requires the use of scientific investigations that target the complexity of our perceptions and how its different ‘ingredients’ mutually influence one another.

time course of crosscategorical affective influence

The empirical data reported here in chapters three to six clearly indicate reciprocal influence between perceiving bodily expressions and contextual stimuli like auditory expressions or background scenes. An important issue concerns when and where in the perceptual system this interaction occurs. The nature and time point of the affective crossmodal and crosscategorical influence can be situated on a continuum with on one extreme interaction in the earliest stages of perceptual processing and on the other end a very late interaction between stimulus categories, implying that both stimulus categories are first processed completely separately and only influence one another during later stages in the sensory processing stream. This latter time point is associated with more cognitive and conscious decisional effects, whereas the early integration assumes perceptual, automatic implicit effects. There is little need to provide evidence for later effects, since the occurrence of crossmodal influence in itself implies that they are situated in the time continuum and positioning near the late extreme is most parsimonious, whereas the argument for early integration needs to be empirically founded. The design of the audiovisual studies reported in chapters four to six followed that of previous studies (B. de Gelder & Vroomen, 2000; Meeren, van Heijnsbergen, & de Gelder, 2005) and does not allow solid conclusions about where on the continuum the crossmodal effects are situated. The use of experimental methodology with high temporal resolution like electro-encephalography or magneto-encephalography are best suited to answer these questions. To date, no such studies have been carried out with bodily expressions, but findings from facial expression research point to early integration of faces and vocal expressions and scenes (B. de Gelder, Vroomen, & Pourtois, 2004; Righart & de Gelder, 2006).

Behavioral experimental techniques exploiting well-documented psychophysical phenomena like audiovisual recalibration aftereffects (e.g. Radeau & Bertelson, 1974) are also useful to investigate the nature of crossmodal bias effects. Audiovisual recalibration has mainly been studied with basic multimodal stimuli like light flashes and sound bursts. For example,

participants are presented with light flashes and beeps that are temporally synchronized, but originating at discordant (incongruent) locations. When the participants are instructed to indicate the location of the auditory stimulus and to ignore the visual stimulus, their responses will be biased towards the location of the visual stimulus. This is known as the ventriloquist illusion (Bertelson, 1999) and this is an immediate effect of intermodal conflict situations. Next to immediate effects, aftereffects occur during unimodal presentation following repeated exposure to incongruent multimodal stimuli. For example, when subjects are exposed to the audiovisual stimuli described above and subsequently presented with an isolated sound burst (without the light flash), the reported location of the auditory stimulus will be biased towards the direction of the light flash, indicating recalibration of auditory localization system. Similar effects have been reported in the temporal domain, where the reported time point shifts towards the time point of the discordant light flash. The occurrence of this kind of after effects argues in favor of early perceptual effects since there is no simultaneous conflicting heteromodal stimulus present. Making use of this paradigm, one could investigate whether recalibration effects also occur in the affective domain. For example, does repeated exposure to an audiovisual neutral body with fearful vocal expression leads to increased 'fearfulness' of the neutral body? So far, no such studies in the affective body-voice domain have been carried out. Maybe the empirical data reported in chapters four to six provide an incentive for future studies addressing this issue.

neuro-anatomy of crosscategorical influences

Besides the time course of intercategorical stimulus integration, another issue concerns the location at the brain level where the reciprocal influence occurs. Again, we have to turn to data from facial expressions to infer hypothesis about whole body expressions. Evidence from multiple face-voice combination fMRI-studies point to involvement of the amygdala in multisensory affective integration (R. J. Dolan, Morris, & de Gelder, 2001; Ethofer et al., 2006) and this finding is supported by anatomical data, showing multisensory projections to amygdala (McDonald, 1998).

Next to the amygdala, recent findings have suggested a role of premotor structures in audiovisual perception. Bimodal mirror neurons have been described in monkey premotor cortex and they fire both when an action is seen or heard (Kohler et al., 2002). The evidence

for the existence of a mirror neuron system in humans is until now indirect (Grèzes, Armony, Rowe, & Passingham, 2003), but recent data are compatible with an audiovisual mirror neuron system in human premotor cortex: the premotor cortex is more more activated by listening to piano excerpts that belong to the motor repertoire of the observer than excerpts for which the observer has no executable skills (Lahav, Saltzman, & Schlaug, 2007).

Thirdly, the role of the superior temporal sulcus (STS) in audiovisual perception as well as processing of biological motion is well established. Recent audiovisual face-voice emotion perception studies also found superadditive activation for audiovisual stimuli compared to the unimodal conditions in the superior temporal region (Robins, Hunyadi, & Schultz, 2009).

Phylogenetically older brain structures, like the superior colliculus contains multimodal topographic maps of the environment in the visual, auditory and tactile domain. Its function is mostly associated with orienting of attention, but it also has a role in emotion processing (Morris, de Gelder, Weiskrantz, & Dolan, 2001; Panksepp, 1998).

Finally, it has been shown that perception of multisensory events, particularly face-voice combinations also mediates activity in primary sensory areas (Ghazanfar, Maier, Hoffman, & Logothetis, 2005).

It is clear from a first glance at the regions discussed above that crosscategorical stimulus influences are mediated through a complex network of brain structures with feedforward and feedback connections. A helpful approach to start to untangle the mechanisms of perception and emotional perception in particular may be the evolutionary perspective. The visual system evolved from the organism's requirement to distinguish brightness from darkness, so its primary functions were quite primitive. In parallel with functional specialization at the central nervous systems, there has also been specialization at the ocular level, leading to differentiated sensitivity across the visual field. This is specialization has its correlates in the early projections from retinal ganglion cells to the evolutionary older brain structures of the visual system.

The evolutionary approach is also particularly relevant regarding emotion, especially whole body expressions that provide a link with adaptive actions. It is therefore not surprisingly that Darwin himself investigated emotional expressions in man and animals (Darwin, 1872). The model displayed below aims to provide a coarse framework for crosscategorical modulation.

Parvocellular ganglion cells with small receptive fields and sensitivity to color and form are primarily present in the foveal part of the retina, while magnocellular retinal ganglion cells with larger receptive fields and sensitivity to motion are more equally distributed across the retina. This division is reflected further in the visual processing stream by the tectopulvinar and geniculostriate pathway, where the latter is primarily involved in detailed form and color perception, and the former is more involved in motion perception and orienting. It is adaptive for the organism to be able to rapidly orient to moving and emotionally significant (and potentially threatening) stimuli in the foveal visual field but also and perhaps even more importantly in the periferral visual field. This function primarily relies on the tectopulvinar pathway and one of the functions of this processing stream lies in detection of emotional stimuli. The geniculostriate pathway is primarily involved in fine grained analysis of form, shape, colour and texture of stimuli in the central visual field. From the striate cortex, detailed processing of different object classes like faces, bodies or houses occurs in specialized areas of the ventral stream. Amygdalar projections to these occipitotemporal areas modulates the processing of emotional stimuli.

The model outlined below is in principle applicable for perception in general. The emphasis here lies on the emotional information conveyed by stimulus categories. The primary question then is what constitutes ‘emotional’. There are legio definitions of emotion in the literature each emphasizing on or another aspect, like subjective feeling, action readiness or cognition. In line with the above reasoning, we would like to put the emphasis on the evolutionary aspect and define emotions as ‘acute mental and bodily states that facilitate adaptive reactions to present environmental events’.

The term ‘acute’ is included to distinguish emotions from moods, which are associated with a more graduate build-up and longer duration. ‘Mental and bodily states’ refer to the subjective feelings that are part of emotions as well as to the associated behavioral patterns and actions expressed by the organism. The latter part of the definition, ‘present environmental events’, refers to the notion that emotions are reactions to very recent or ongoing events that occur in the direct surroundings of the organism.

In conclusion, the findings discussed here point to reciprocal influences between the many ingredients making up the complex course of our everyday perceptions. Evolution has shaped

our brain and behavior with ‘special’ sensitivity for perception of emotional expressions conveyed by faces, voices bodies and scenes.

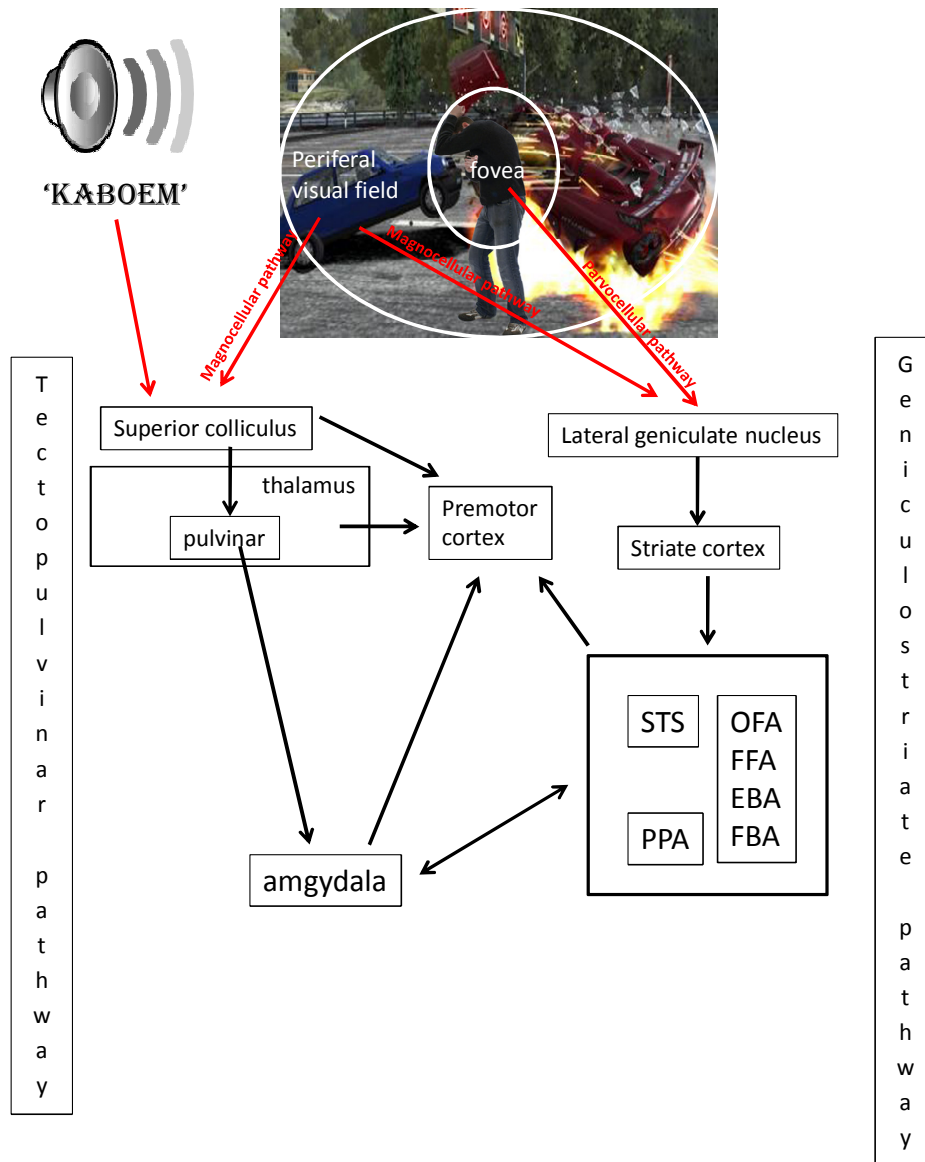


Figure 1. Neural pathways for processing auditory and visual emotional information

References

- Adolphs, R. (2002). Neural systems for recognizing emotion. *Curr Opin Neurobiol*, 12(2), 169-177.
- Argyle, M. (1988). *Bodily communication*. London: Methuen.
- Bertelson, P. (1999). Ventriloquism: A case of crossmodal perceptual grouping. In G. Aschersleben & T. Bachmann (Eds.), *Cognitive contributions to the perception of spatial and temporal events. Advances in psychology*, 129 (pp. 347-362). Amsterdam, Netherlands: North-Holland/Elsevier Science Publishers.
- Cutting, J. E., & Kozlowski, L. T. (1977). Recognizing friends by their walk: Gait perception without familiarity cues. *Bulletin of the Psychonomic Society*, 9, 353-356.
- Darwin, C. (1872). *The expression of the emotions in man and animals*. London: John Murray.
- de Gelder, B. (2006). Towards the neurobiology of emotional body language. *Nature Reviews Neuroscience*, 7(3), 242-249.
- de Gelder, B. (2009). Why bodies? Twelve reasons for including bodily expressions in affective neuroscience. *Proc R Soc Lond B Biol Sci*, (in press).
- de Gelder, B., Snyder, J., Greve, D., Gerard, G., & Hadjikhani, N. (2004). Fear fosters flight: A mechanism for fear contagion when perceiving emotion expressed by a whole body. *Proc Natl Acad Sci U S A*, 101(47), 16701-16706.
- de Gelder, B., Van den Stock, J., Meeren, H. K., Sinke, C. B., Kret, M. E., & Tamietto, M. (2009). Standing up for the body. Recent progress in uncovering the networks involved in processing bodies and bodily expressions. *Neuroscience and Biobehavioral Reviews*.
- de Gelder, B., & Vroomen, J. (2000). The perception of emotions by ear and by eye. *Cognition and Emotion*, 14(3), 289-311.
- de Gelder, B., Vroomen, J., & Pourtois, G. (2004). Multisensory perception of emotion, its time course and its neural basis. In G. Calvert, C. Spence & B. E. Stein (Eds.), *Handbook of multisensory processes*. (pp. 581-596). Cambridge, MA: MIT.
- Dolan, R. J. (2002). Emotion, Cognition and Behavior. *Science*, 8(298), 1191-1194.
- Dolan, R. J., Morris, J. S., & de Gelder, B. (2001). Crossmodal binding of fear in voice and face. *Proc Natl Acad Sci U S A*, 98(17), 10006-10010.
- Ekman, P., & Friesen, W. V. (1976). *Pictures of facial affects*. Palo Alto: Consulting Psychologists Press.
- Ethofer, T., Anders, S., Erb, M., Droll, C., Royen, L., Saur, R., et al. (2006). Impact of voice on emotional judgment of faces: an event-related fMRI study. *Hum Brain Mapp*, 27(9), 707-714.
- Gauthier, I., Skudlarski, P., Gore, J. C., & Anderson, A. W. (2000). Expertise for cars and birds recruits brain areas involved in face recognition. *Nature Neuroscience*, 3(2), 191-197.

- Ghazanfar, A. A., Maier, J. X., Hoffman, K. L., & Logothetis, N. K. (2005). Multisensory integration of dynamic faces and voices in rhesus monkey auditory cortex. *J Neurosci*, 25(20), 5004-5012.
- Grèzes, J., Armony, J. L., Rowe, J., & Passingham, R. E. (2003). Activations related to "mirror" and "canonical" neurones in the human brain: an fMRI study. *Neuroimage*, 18(4), 928-937.
- Grèzes, J., Pichon, S., & de Gelder, B. (2007). Perceiving fear in dynamic body expressions. *Neuroimage*, 35(2), 959-967.
- Kohler, E., Keysers, C., Umiltà, M. A., Fogassi, L., Gallese, V., & Rizzolatti, G. (2002). Hearing sounds, understanding actions: action representation in mirror neurons. *Science*, 297(5582), 846-848.
- Lahav, A., Saltzman, E., & Schlaug, G. (2007). Action representation of sound: audiomotor recognition network while listening to newly acquired actions. *J Neurosci*, 27(2), 308-314.
- McDonald, A. J. (1998). Cortical pathways to the mammalian amygdala. *Prog Neurobiol*, 55(3), 257-332.
- Meeren, H. K., van Heijnsbergen, C. C., & de Gelder, B. (2005). Rapid perceptual integration of facial expression and emotional body language. *Proc Natl Acad Sci U S A*, 102(45), 16518-16523.
- Morris, J. S., de Gelder, B., Weiskrantz, L., & Dolan, R. J. (2001). Differential extrageniculostriate and amygdala responses to presentation of emotional faces in a cortically blind field. *Brain*, 124(Pt 6), 1241-1252.
- Panksepp, J. (1998). *Affective neuroscience: The foundation of human and animal emotions*. New York: Oxford University Press.
- Peelen, M. V., & Downing, P. E. (2007). The neural basis of visual body perception. *Nat Rev Neurosci*, 8(8), 636-648.
- Phelps, E. A., & LeDoux, J. E. (2005). Contributions of the amygdala to emotion processing: from animal models to human behavior. *Neuron*, 48(2), 175-187.
- Pichon, S., de Gelder, B., & Grèzes, J. (2008). Emotional modulation of visual and motor areas by dynamic body expressions of anger. *Social Neuroscience*, 3(3-4), 199-212.
- Preston, S. D., & de Waal, F. B. (2002). Empathy: Its ultimate and proximate bases. *Behav Brain Sci*, 25(1), 1-20; discussion 20-71.
- Radeau, M., & Bertelson, P. (1974). The after-effects of ventriloquism. *Q J Exp Psychol*, 26(1), 63-71.
- Righart, R., & de Gelder, B. (2006). Context influences early perceptual analysis of faces--an electrophysiological study. *Cereb Cortex*, 16(9), 1249-1257.
- Rizzolatti, G., & Craighero, L. (2004). The mirror-neuron system. *Annu Rev Neurosci*, 27, 169-192.
- Robins, D. L., Hunyadi, E., & Schultz, R. T. (2009). Superior temporal activation in response to dynamic audio-visual emotional cues. *Brain Cogn*, 69(2), 269-278.
- Schmidt, K. L., & Cohn, J. F. (2001). Human facial expressions as adaptations: Evolutionary questions in facial expression research. *Am J Phys Anthropol*, Suppl 33, 3-24.

- Sprengelmeyer, R., Young, A. W., Schroeder, U., Grossenbacher, P. G., Federlein, J., Buttner, T., et al. (1999). Knowing no fear. *Proc Biol Sci*, 266(1437), 2451-2456.
- Van den Stock, J., Righart, R., & de Gelder, B. (2007). Body expressions influence recognition of emotions in the face and voice. *Emotion*, 7(3), 487-494.

Acknowledgements

Working with Bea de Gelder has been a unique experience in too many ways to sum up. Despite her contagious professional passion, relentless scientific insight and occasional materialistic hedonism, I simply enjoyed just working with her. Bea, that sweltering day in the summer of 2003 I set foot in your office can only be a flashbulb memory considering its detailed consolidation in my mind. I started my PhD one year later and work has never been so much fun.

Many others contributed to the pleasant and inspiring atmosphere in the lab and beyond: Jeroen, Hanneke, Ilja, Mirjam, Jean, Geert, Ton, Maurice, Sabine, Mariska, Bernard, Charlotte, Wim, Rhutgher, Margriet, Carlijn, Karen, Martijn, Liselotte, Viona, Tessa, Stijn, Siebe, Renee, Lieven, Pieter, Mehrdad, Lisanne, Ankie, Johan, Gert, Ivar, Julie, Swann, Jane, Wobbe, Marco and all the others I shamefully but unintentionally left out.

Further, I am grateful for enlightening discussions and learning opportunities with researchers and clinicians beyond Tilburg University, not in the least Mathieu Vandenbulcke, whose supportive attitude is only exceeded by his professional integrity. I would also like to thank Rik Vandeberghe and the entire cognitive neurology lab at Leuven, Kathleen Porke, Luc Van de Ven, Wim Van der Elst, Rainer Goebel, Bettina Sorger, Sven Gijssen, Ron Peeters, Gert Cypers and the other neurologists from the OLV Aalst, Wim Vanduffel, Wouter Depuyt, Sabine Deprez, Katrien Vanderperren, Nikolay Novitski, Koen Nelissen, Qi Zhu and all the others who directly or indirectly and willingly or unwillingly contributed to this thesis.

Unfortunately, there are at least three people to whom I can only express my gratitude posthumously: Petjen Beir, who is responsible for the earliest stimulation of my cognitive functioning, Erik Vandenbussche and Paul Bertelson.

Finally, the support from my family can hardly be underestimated. Patty, although work is as fulfilling as can be, nothing tops coming home to Toos and you.